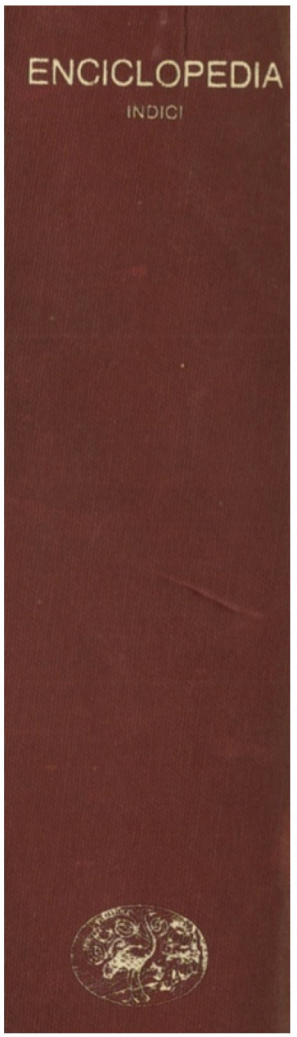


ENCICLOPEDIA EINAUDI [1982]

FISICA

Paolo Caldirola e Giulio Casati - **FISICA** pag.4

Giancarlo Ramunni - **ATOMO E MOLECOLA** pag.10

Yehuda Elkana e Yemima Ben-Menahem
- **CONSERVAZIONE/INVARIANZA** pag.26
ENTROPIA pag.45

Bernard d'Espagnat - **FISICA** pag.62

Antonio Sparzani - **FORZA/CAMPO** pag.82

Carlo Ferrari - **MOTO** pag.93

Michael O'Hanlon - **PARTICELLA** pag.143

Bruno Bertotti - **PLASMA** pag.164

Yvonne Choquet-Bruhat - **PROPAGAZIONE** pag.177

Antonio Sparzani - **QUANTI** pag.195

Tullio Regge - **RELATIVITÁ** pag.214

Francesco Guerra - **REVERSIBILITÁ/IRREVERSIBILITÁ** pag.249

Luigi Accardi -**STATO FISICO** pag.270

ambiguità
competenza/esecuzione
fonetica
grammatica
lessico
lingua
lingua/parola
linguaggio
metrica
semantica
senso/significato
traduzione
universali/particolari

concetto
esistenza
essere
fenomeno
forma
idea
proposizione e giudizio

analogia e metafora
argomentazione
interpretazione

astratto/concreto
dialettica
identità/differenza
mediazione
opposizione/contraddizione
qualità/quantità
totalità
uno/molti

atti linguistici
dicibile/indicibile
enunciazione
presupposizione e allusione
referente

comunicazione
errore
informazione

decisione
distribuzione statistica
giochi
induzione statistica
probabilità
rappresentazione statistica
teoria/pratica

dato
etica
filosofia/filosofie
ragione
razionale/irrazionale
soggetto/oggetto
uguaglianza
valori
vero/falso
volontà

catastrofi
ciclo
evento
periodizzazione
tempo/temporalità

antico/moderno
calendario
decadenza
escatologia
età mitiche
genesi
passato/presente
progresso/reazione
storia

curve e superfici
geometria e topologia
invariante

caos/cosmo
infinito
macrocosmo/microcosmo
mondo
natura
osservazione
reale
unità

alchimia
astrologia
cabala
elementi
esoterico/essoterico

atlante
collezione
documento/monumento
fossile
memoria
rovina/restauro

armi
frontiera
guerra
imperi
nazione
tattica/strategia

deduzione/prova
equivalenza
formalizzazione
logica
possibilità/necessità
referenza/verità
ricorsività

differenziale
funzioni
infinitesimale
locale/globale
sistemi di riferimento
stabilità/instabilità
variazione

analisi/sintesi
anticipazione
ipotesi
matematiche
metodo

funzione
misura
modello
struttura
teoria/modello

alienazione
coscienza/autocoscienza
immaginazione sociale
pace
servo/signore
uomo
utopia

demagogia
discriminazione
repressione
terrore
tolleranza/intolleranza
tortura
violenza

centrato/acentrato
combinatoria
grafo
labirinto
rete

applicazioni
assioma/postulato
continuo/discreto
dipendenza/indipendenza
divisibilità
dualità
insieme
razionale/algebrico/trascendente
simmetria
strutture matematiche
trasformazioni naturali / categorie

abaco
algoritmo
approssimazione
calcolo
numero
zero

caso/probabilità
causa/effetto
certezza/dubbio
coerenza
convenzione
determinato/indeterminato
empiria/esperienza
esperimento
legge
libertà/necessità
metafisica
naturale/artificiale
operatività
paradigma
previsione e possibilità
riduzione
ripetizione
scienza
spiegazione
verificabilità/falsificabilità

categorie/categorizzazione
conoscenza
coppie filosofiche
disciplina/discipline
enciclopedia
innovazione/scoperta
insegnamento
invenzione
rappresentazione
ricerca
sistematica e classificazione

controllo/retroazione
energia
equilibrio/squilibrio
interazione
ordine/disordine
organizzazione
semplice/complesso
sistema
soglia
vincolo

analogico/digitale
automa
intelligenza artificiale
macchina
programma
simulazione
strumento

cervello
comportamento
e condizionamento
controllo sociale
emozione/motivazione
mente

apprendimento
autoregolazione/equilibrazione
cognizione
induzione/deduzione
innato/acquisito
istinto
operazioni
percezione
quoziente intellettuale

astronomia
cosmologie
gravitazione
luce
materia
spazio-tempo

atomo e molecola
conservazione/invarianza
entropia
fisica
forza/campo
moto
particella
plasma
propagazione
quanti
relatività

atmosfera
litosfera
oceani
pianeti
sole
universo

reversibilità/irreversibilità
stato fisico

adattamento
evoluzione
mutazione/selezione
polimorfismo
specie

cellula
differenziamento
immunità
individualità biologica
integrazione
invecchiamento
organismo
regolazione
sviluppo e morfogenesi

catalisi
macromolecole
metabolismo
omeostasi
organico/inorganico
osmosi
vita

eredità
gene
genotipo/fenotipo
razza
sangue

allegoria
codice
immagine
metafora
segno
significato
simbolo

avanguardia
classico
critica
filologia
letteratura
maniera
poetica
retorica

bello/brutto
creatività
espressione
fantastico
gusto
imitazione
immaginazione
progetto
riproduzione/riproducibilità
sensibilità
spazialità

anthropos
cultura/culture
etnocentrismi
natura/cultura

alfabeto
ascolto
gesto
lettura
luogo comune
orale/scritto
parola
ritmo
scrittura
voce

discorso
finzione
generi
narrazione/narratività
stile
tema/motivo
testo

arti
artigianato
artista
attribuzione
oggetto
produzione artistica

acculturazione
civiltà
futuro
selvaggio/barbaro/civilizzato

armonia
melodia
ritmica/metrica
scala
suono/rumore
tonale/atonale

colore
disegno/progetto
visione

escrementi
fertilità
nascita
sensi
sessualità
vecchiaia
vita/morte

abbigliamento
canto
corpo
danza
maschera
moda
ornamento
scena

educazione
generazioni
infanzia
morte

coltivazione
cultura materiale
industria rurale
materiali
prodotti

credenze
dialetto
enigma
fiaba
mostro
popolare
proverbi
tradizioni

amore
desiderio
eros
isteria
pulsione
soma/psiche
sonno/sogno

angoscia/colpa
castrazione e complesso
censura
identificazione e transfert
inconscio
nevrosi/psicosi
piacere

clinica
cura/normalizzazione
esclusione/integrazione
farmaco/droga
follia/delirio
medicina/medicalizzazione
normale/anormale
salute/malattia
sintomo/diagnosi

fuoco
homo
mano/manufatto
tecnica
utensile

cannibalismo
dèi
divino
eroi
iniziazione
magia
messia
millennio
persona
puro/impuro
religione
sogno/visione
stregoneria

ateo
chierico/laico
chiesa
diavolo
eresia
libertino
libro
peccato
sacro/profano
santità

demoni
divinazione
mito/rito
mythos/logos
origini

agonismo
cerimoniale
festa
feticcio
gioco
lutto
regalità
rito

casta
donna
endogamia/esogamia
famiglia
incesto
maschile/femminile
matrimonio
parentela
totem
uomo/donna

alimentazione
animale
cucina
domesticamento
fame
vegetale

caccia/raccolta
dono
eccedente
pastorizia
primitivo
reciprocità/ridistribuzione

borghesi/borghesia
burocrazia
classi
contadini
ideologia
masse
proletariato
rivoluzione

economia
formazione economico-sociale
lavoro
modo di produzione
proprietà
riproduzione
transizione

consenso/dissenso
egemonia/dittatura
intellettuali
libertà
maggioranza/minoranza
partiti
politica

abbondanza/scarsità
bisogno
consumo
imposta
lusso
oro e argento
pesi e misure
produzione/distribuzione
ricchezza
scambio
spreco

amministrazione
comunità
conflitto
élite
gergo
gruppo
marginalità
opinione
povertà
propaganda
ruolo/status
socializzazione
società
spazio sociale

accumulazione
capitale
crisi
distribuzione
fabbrica
gestione
imperialismo
impresa
mercato
merce
moneta
pianificazione
profitto
rendita
salario
utilità
valore/plusvalore

consuetudine
diritto
giustizia
istituzioni
responsabilità

costituzione
democrazia/dittatura
norma
patto
potere
potere/autorità
pubblico/privato
società civile
stato

abitazione
acqua
ambiente
città
clima
ecumene
insediamento
migrazione
paesaggio
popolazione
regione
risorse
suolo
terra
territorio
villaggio

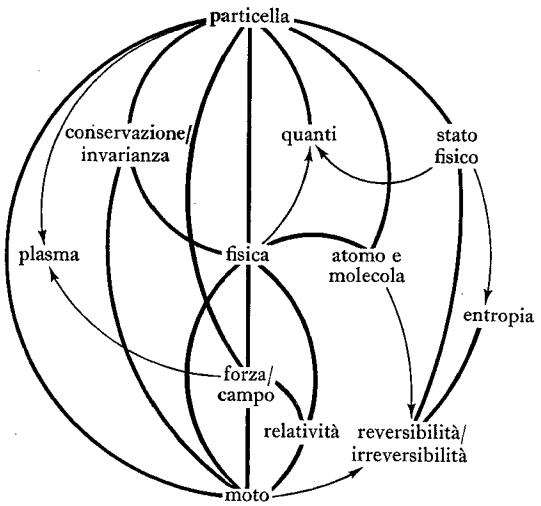
agricoltura
città/campagna
colonie
commercio
industria
spazio economico
sviluppo/sottosviluppo

Fisica

	anthropos	argomentazione	arti	calcolo	capitale	cervello	civiltà	classi	cognizione	combinatoria	comunicazione	concetto	conoscenza	corpo	creatività	cultura mat.	dialettica	domesticam.	eccedente	enunciazione	eredità	esoterico/essoter.	evoluzione	filosofia/filosofie	**fisica**	generazioni	geometria e topol.	giustizia	guerra	homo	inconscio	letteratura	linguaggio	locale/globale	logica	macchina	materia
atomo e molecola	·	·	2	·	·	2	·	·	3	4	5	4	5	·	2	·	6	·	3	·	6	4	5	2	9	·	·	·	·	3	·	·	·	4	4	·	8
conservazione/invar.	2	·	·	·	·	2	·	·	3	2	5	2	4	·	·	·	3	1	·	2	·	·	·	·	8	3	4	·	·	·	·	·	·	5	·	2	6
entropia	3	·	2	·	3	·	4	2	·	·	7	·	·	2	·	·	2	·	·	·	·	·	5	·	6	·	·	·	·	·	·	·	·	·	·	·	4
fisica	5	·	5	4	·	5	·	3	3	3	5	6	8	·	5	·	7	·	·	·	4	4	4	6	9	·	3	·	1	3	1	3	3	6	7	5	8
forza/campo	3	·	1	3	·	2	1	2	3	1	5	5	4	2	·	4	6	·	3	·	·	3	·	3	9	·	7	2	3	3	3	2	2	6	4	6	9
moto	·	·	·	3	2	2	·	·	2	·	4	4	5	·	·	·	7	·	·	·	·	·	5	·	9	·	4	·	·	·	·	·	·	6	4	4	9
particella	·	·	·	·	·	2	·	·	1	3	4	2	4	·	2	·	5	·	·	·	3	3	3	1	9	·	·	·	·	·	·	·	·	5	5	·	9
plasma	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	2	·	·	·	·	·	·	·	·	·	·	·	·
propagazione	·	·	·	4	·	·	·	·	·	·	·	3	2	·	·	·	2	·	·	·	·	·	·	·	4	·	·	·	·	·	·	·	3	3	·	·	5
quanti	·	·	·	3	·	·	·	·	·	·	·	4	·	·	2	·	3	·	·	·	3	3	·	·	7	·	·	·	·	·	·	·	·	3	4	·	6
relatività	·	·	·	·	·	·	·	·	·	·	3	4	4	·	3	·	6	·	·	·	·	·	·	·	8	·	5	·	·	·	·	·	·	6	6	·	8
reversibilità/irreversib.	·	·	·	·	·	4	·	·	5	·	·	·	·	·	4	·	4	·	·	·	·	·	·	2	6	·	·	·	·	·	·	·	·	3	·	·	2
stato fisico	·	·	·	·	·	7	·	·	5	7	6	5	·	·	·	·	5	·	·	2	·	3	6	1	9	·	·	·	·	4	·	·	6	1	4	6	6

	memoria	metodo	modo di produz.	mythos/logos	natura	normale/anormale	orale/scritto	organico/inorgan.	organismo	parentela	politica	probabilità	produzione/distrib.	regione	religione	rito	sacro/profano	segno	sistema	società	soma/psiche	spiegazione	stato	storia	strutture matem.	sviluppo/sottosvil.	tempo/temporalità	teoria/modello	testo	tolleranza/intoll.	tonale/atonale	tradizioni	universo	uomo	visione	vita/morte
atomo e molecola	·	4	·	·	5	2	3	9	8	·	·	3	2	4	·	·	2	4	7	·	2	6	·	3	7	·	5	5	·	·	·	·	7	2	·	·
conservazione/invar.	1	3	3	·	3	·	2	3	5	·	·	·	3	3	·	·	·	·	7	2	·	6	·	·	6	·	5	5	·	2	·	1	5	2	·	2
entropia	3	3	·	·	·	·	3	6	·	·	2	2	2	·	·	·	·	·	7	·	4	3	2	·	4	·	5	·	·	·	·	·	4	·	3	·
fisica	2	6	3	2	7	4	3	6	6	·	3	3	4	4	3	·	5	4	8	2	3	8	·	4	8	3	7	7	·	2	·	3	7	4	3	·
forza/campo	·	5	3	2	5	·	2	5	5	2	·	4	4	2	5	·	4	2	7	4	5	7	·	3	8	3	5	6	2	·	·	2	7	2	2	·
moto	·	5	·	·	6	2	1	3	·	·	·	3	·	·	·	·	2	·	7	·	·	7	·	2	7	·	5	6	·	·	·	2	8	·	4	·
particella	·	5	·	·	6	·	·	5	5	·	·	3	·	5	·	·	·	·	6	·	·	6	·	2	8	·	3	2	1	·	·	·	7	·	·	·
plasma	·	·	·	·	·	·	·	5	6	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	6	·	·	·
propagazione	·	2	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	4	·	2	·	·	·	3	·	·	·	·	·	·	·	4	·	·	·
quanti	·	·	·	·	·	·	·	4	·	·	·	3	·	·	·	·	·	·	4	·	·	3	·	3	7	·	3	4	·	·	·	·	2	·	·	·
relatività	·	5	·	1	4	·	3	·	3	·	·	3	1	·	3	·	·	·	5	·	·	7	·	·	7	·	5	6	·	·	·	·	5	·	·	·
reversibilità/irreversib.	·	·	·	2	·	3	·	4	·	·	·	·	·	·	·	·	·	4	7	·	·	·	·	2	3	·	6	3	2	·	·	·	·	2	3	2
stato fisico	·	3	·	·	4	6	2	7	6	·	·	1	·	3	·	·	·	·	9	3	4	8	·	5	5	·	6	4	·	·	·	·	7	4	·	·

	relatività	moto	forza/campo	particella	fisica	conservazione/invarianza	atomo e molecola	quanti	stato fisico	reversibilità/irreversibilità	entropia	plasma	propagazione
relatività	·	5	4	3	5	4	2	4	4	2	·	·	3
moto	7	·	5	6	3	6	3	5	·	6	4	5	5
forza/campo	7	6	·	6	6	5	4	4	3	5	2	6	5
particella	5	6	5	·	6	6	6	5	5	4	4	6	3
fisica	5	6	5	6	·	6	5	6	4	4	5	5	4
conservazione/invarianza	3	4	3	6	3	·	2	4	3	4	4	4	3
atomo e molecola	3	4	3	6	6	4	·	5	·	6	5	5	·
quanti	4	3	2	5	·	4	4	·	·	·	·	·	·
stato fisico	·	5	4	6	5	4	4	6	·	6	6	5	4
reversibilità/irreversibilità	·	·	·	·	2	·	·	·	2	·	5	3	·
entropia	·	2	·	·	·	2	·	3	·	5	·	2	1
plasma	·	·	·	·	·	·	·	·	·	1	·	·	·
propagazione	·	·	·	·	·	·	·	1	·	·	·	3	·



Fisica

Atomo e molecola, Conservazione/invarianza, Entropia, Fisica, Forza/campo, Moto, Particella, Plasma, Propagazione, Quanti, Relatività, Reversibilità/irreversibilità, Stato fisico

Il problema centrale, di fondo, della ♦fisica♦ non è molto cambiato da quello che fu il fondamentale contributo degli antichi pensatori greci; come non sono cambiate le difficoltà. Ridotto ai suoi termini essenziali, il problema ultimo rimane ancora quello d'interpretare il comportamento degli oggetti che ci circondano a partire dal ♦moto♦ dei costituenti «elementari». È noto che la varietà delle forme e delle proprietà sotto cui si presenta la materia è dovuta alla diversa combinazione e proporzione degli elementi fondamentali che la costituiscono. Il loro moto determina proprietà come il colore, il sapore, l'odore, ecc. Ma esistono i costituenti «elementari» della materia? Quale il loro moto e quali le cause che lo determinano? Esiste una forza fondamentale della natura all'origine di tutti questi moti? Non solo, ma esiste un legame tra i costituenti elementari della materia e le forze che ne determinano il moto, oppure sono da considerarsi come entità indipendenti?

La prima grande autorevole risposta a questi interrogativi è stata data, come piú volte ricordato, da Aristotele. Egli introdusse il concetto di moto naturale, che è quello che compete ai quattro elementi fondamentali – terra, acqua, aria, fuoco – di cui è costituita tutta la materia. Ciò che presiede a tutti i movimenti è l'esistenza di una causa finale per cui tutti gli elementi tendono a raggiungere il loro luogo naturale. Ne segue che una pietra, composta in massima parte dall'elemento terra, tende a cadere attraverso l'aria e l'acqua per congiungersi con la Terra, suo luogo naturale. Allo stesso modo il fuoco sale, essendo il suo luogo naturale posto nella posizione piú elevata; ecc. I corpi celesti invece, perfetti, sono costituiti da una quinta sostanza (quintessenza), immutabile ed incorruttibile, e di conseguenza non possono che muoversi nel modo piú perfetto possibile, quello del moto circolare uniforme.

La grande costruzione di Aristotele doveva crollare circa duemila anni dopo, ad opera di Galileo e di Newton. Quest'ultimo diede le tre leggi fondamentali della dinamica che regolavano non solo il moto di tutti gli oggetti terrestri ma anche dei corpi celesti. I complessi meccanismi del sistema tolemaico – epicicli, equanti, eccentrici, ecc. – vennero sostituiti da un'unica grande legge: la legge della gravitazione universale. Ma, come osserva Thom, con la legge di gravitazione universale Newton calcolava tutto e non spiegava nulla e oggi non ci sono meno ragioni di stupirsi della caduta della mela di quante non ne avesse Newton. Tuttavia la necessità di sapere dove sarebbe andata a cadere esattamente una palla di cannone ebbe il sopravvento sull'autorità e sull'opinione di Aristotele: la storia diede ragione a Newton. Certo rimaneva aperto il problema interpretativo al di là del nome «che piú singulare e proprio gli abbiamo assegnato di gravità» (Galileo). Rimaneva da «spiegare» tale forza e come essa

si potesse trasmettere nello spazio da un pianeta all'altro senza il supporto di un mezzo interposto. Non solo, ma un esame piú attento e critico crea ulteriori dubbi. La presenza di «azioni a distanza» come quelle dovute all'interazione gravitazionale toglie significato al concetto primitivo ed intuitivo di forza che si esercita mediante contatto e getta delle ombre sulla fondatezza delle stesse leggi della dinamica.

La grande varietà di forze che noi sperimentiamo a livello macroscopico – la forza del vento, le forze muscolari, il peso dei corpi, ecc. – si riduce a due sole forze fondamentali, la forza gravitazionale e la forza elettromagnetica. La forza gravitazionale è molto debole, circa 10^{38} volte piú piccola di quella elettromagnetica: basta strofinare un pezzo d'ambra per produrre una forza elettrica capace di attrarre dei pezzettini di carta e di opporsi perciò alla forza gravitazionale prodotta dall'intera massa terrestre. Queste erano le sole forze note fino al secolo scorso ed era naturale il tentativo di farle derivare da un'unica forza fondamentale. Questo sforzo di unificazione, di descrivere i fenomeni fisici attraverso qualcosa di elementare o di «semplice», è alla base della scoperta scientifica. E l'idea che dietro l'apparente diversità e complessità degli oggetti che ci circondano sia nascosta una sostanziale semplicità di fondo, ha ispirato il pensiero scientifico fin dai tempi piú antichi, da quando Democrito nel v secolo a. C. avanzò l'ipotesi che la materia fosse composta da innumerevoli particelle indivisibili (atomi) in movimento all'interno dei corpi le cui diverse proprietà ed apparenze erano dovute alle diverse combinazioni degli atomi. Del resto, «spiegare» un certo fenomeno cosa vuol dire se non cercare di ricondurre il comportamento a poche semplici leggi? Le equazioni fondamentali della dinamica, ad esempio, descrivono ugualmente bene la caduta di un sasso, il moto di una freccia o il percorso di una cometa. Che questa riduzione dei fenomeni piú disparati a poche semplici leggi sia possibile è una straordinaria e meravigliosa proprietà della Natura.

Un passo decisivo in avanti rispetto alla cosmologia newtoniana è stato fatto ad opera di Maxwell con l'introduzione del concetto di campo e della dicotomia ♦forza/campo♦: esso, assieme al concetto di energia, ha permesso di capire come le forze possono agire a distanza senza la necessità di un mezzo che le trasmetta. Partendo dai lavori di Faraday sulle forze elettriche e magnetiche, Maxwell introdusse questa quantità invisibile chiamata campo e riuscí a scrivere le equazioni a cui obbedivano i campi elettrici e magnetici, fornendo cosí una teoria unificata del campo elettromagnetico. (Il processo di unificazione dei due campi verrà completato con la teoria della relatività di Einstein).

Ma il piú grande successo delle equazioni di Maxwell è stato l'aver fatto comprendere che l'ottica non era null'altro che un capitolo dell'elettromagnetismo e l'aver previsto l'esistenza di onde elettromagnetiche che si propagano nel vuoto senza alcun supporto materiale trasportando, con sé, energia; da tali onde dipende l'esistenza della vita stessa sulla Terra.

Il grande passo in avanti compiuto con Maxwell nella descrizione del mondo fisico è stato possibile grazie a uno sforzo di astrazione e a un uso ampio dello strumento matematico. Queste caratteristiche sono destinate ad accentuarsi sem-

pre piú in seguito. Il progressivo allontanamento tra la comprensione del fenomeno fisico e l'intuizione e sensibilità quotidiana raggiunge un livello straordinariamente elevato nella teoria della ◆relatività◆ e nella meccanica quantistica; di pari passo il formalismo matematico diventa piú pesante e le nuove teorie fisiche fanno uso di teorie matematiche sempre piú astratte e complesse. Molti dei progressi della ◆fisica◆ degli ultimi decenni si fondano ad esempio sulla nozione matematica di gruppo, anche perché i teoremi di ◆conservazione◆ che caratterizzano i processi fisici appaiono legati alla ◆invarianza◆ delle leggi fisiche rispetto a particolari trasformazioni. Nel contempo però lo strumento matematico ha anche una notevole funzione unificante, dal momento che il suo crescente impiego contribuisce in maniera determinante a far sí che la tradizionale divisione della fisica nelle diverse branche vada perdendo sempre piú significato.

La speranza di unificare tutte le forze conosciute, che fino al secolo scorso erano solo quella gravitazionale e quella elettromagnetica, che era stata una delle piú salde convinzioni di Einstein, sembrò svanire definitivamente quando si scoprí l'esistenza di una terza forza con caratteristiche diverse dalle prime due e che agiva a livello nucleare. Come è noto, la materia è costituita da atomi di dimensioni molto piccole (circa 10^{-8} cm e con peso di circa 10^{-24} g). Gli atomi però non sono oggetti solidi e indivisibili, ma sono composti a loro volta da un nucleo, carico positivamente, in cui è concentrata quasi tutta la massa, e da particelle molto piú piccole, gli elettroni, carichi negativamente e che possono essere pensati come ruotanti attorno al nucleo molto velocemente (circa mille chilometri al secondo). Secondo questo primo grossolano modello proposto da Rutherford, l'atomo assomiglia al nostro sistema solare in miniatura: il nucleo carico positivamente attrae gli elettroni, i quali tuttavia non cadono sul nucleo grazie alla forza centrifuga.

Una spiegazione piú completa della stabilità degli atomi fu ottenuta però solo successivamente nel decennio 1920-30, con l'avvento della meccanica quantistica.

Questa teoria trasse le sue mosse da alcune osservazioni precedenti, dovute essenzialmente a Niels Bohr, secondo le quali gli elettroni atomici si muovono solo su ben determinate orbite fisse in corrispondenza di una ben determinata energia. Inoltre gli scambi di energia fra un elettrone orbitante e il campo della radiazione elettromagnetica non avvengono in modo continuo ma solamente per multipli interi di certe quantità elementari che dipendono dalla frequenza della radiazione e che sono pertanto detti ◆quanti◆. Questi quanti, o fotoni, hanno le caratteristiche di una ◆particella◆ con massa nulla che ha velocità di ◆propagazione◆ pari alla velocità della luce. La meccanica quantistica, costruita dal lavoro di Broglie, Schrödinger, Heisenberg, Dirac e altri, pervenne a una descrizione compatta delle proprietà sia della materia sia della radiazione risolvendo l'apparente contraddizione del dualismo ondulatorio-corpuscolare che queste due entità sembrano inevitabilmente esibire. La meccanica quantistica, che costituisce una delle maggiori costruzioni del pensiero scientifico di tutti i tempi, non solo ha modificato la nostra comprensione della «struttura» della materia e della radiazione, ma ha portato a una concezione della natura

profondamente diversa dalla concezione intuitiva classica che si era andata formando attraverso l'opera di Newton e di Maxwell. Uno degli aspetti piú caratteristici di questa nuova costruzione teoretica è senza dubbio la sostituzione del rigido determinismo della meccanica newtoniana con una descrizione basata sul concetto di probabilità e la conseguente modifica della definizione di ◆stato fisico◆ di un sistema.

Tornando al problema fondamentale della struttura della materia, fu ben presto riconosciuto che la parte piú interna dell'atomo, il nucleo, è costituito a sua volta da protoni e neutroni. I protoni sono particelle che pesano circa 1836 volte l'elettrone e sono dotati di una carica positiva uguale e contraria a quella dell'elettrone. I neutroni sono particelle elettricamente neutre che pesano quasi esattamente come i protoni. Poiché i protoni sono tutti carichi positivamente, la forza elettrica con la quale si respingono dovrebbe portare a far esplodere il nucleo: infatti le forze gravitazionali, attrattive, sono troppo deboli per bilanciare la repulsione elettrica; deve perciò esistere una nuova forza della natura – la forza nucleare – che tiene uniti protoni e neutroni dentro il nucleo. Gli esperimenti hanno mostrato che la forza nucleare è molto grande e ha un raggio d'azione molto piccolo (circa 10^{-14} m) e questo spiega la ragione per cui questa forza non è osservabile nei fenomeni che si presentano al livello della nostra vita quotidiana. Il numero delle forze fondamentali della natura sale cosí da due a tre. Le speranze precedenti di unificare le due forze fino ad allora conosciute, la gravitazionale e l'elettromagnetica, svaní appunto quando si scoprí questa terza forza cosí diversa dalle altre due.

La stabilità dei nuclei nasce dunque da un equilibrio tra le forze elettriche repulsive e quelle nucleari attrattive e il fenomeno della radioattività consiste proprio in una emissione di particelle dal nucleo nel tentativo, da parte delle particelle costituenti, di mantenere l'equilibrio fra queste due forze.

Lo studio delle forze nucleari e lo sviluppo della teoria matematica concernente soprattutto l'unione, che necessariamente deve esistere, fra la teoria quantistica e la relatività ha mostrato che la *natura delle forze è inseparabile dalla struttura della materia.* Si è scoperto infatti che non solo le proprietà delle particelle dipendono dalle forze della natura ma che le forze stesse si trasmettono mediante particelle le quali sono soggette ad altre forze e cosí via.

Gli sviluppi della teoria hanno anche portato Dirac a proporre l'ardita ipotesi che il «vuoto» non sia, in realtà, per niente vuoto ma riempito da un mare invisibile di particelle con energia negativa. Quando una di queste particelle passa dallo stato di energia negativa in cui si trova a uno stato di energia positiva, allora essa diventa visibile: cioè i «buchi» lasciati da queste particelle invisibili quando abbandonano il mare negativo sono equivalenti alla presenza di particelle visibili: queste ultime particelle costituiscono l'antimateria. Esse, previste dapprima dalla teoria, sono poi state effettivamente osservate; il 2 agosto 1932 il fisico americano Carl Anderson ha osservato per la prima volta l'antiparticella dell'elettrone – il positone –, cosí denominato in quanto ha carica positiva, mentre nel 1955 e 1956 sono stati osservati gli antiprotoni e gli antineutroni. Il nome di antimateria deriva dal fatto che quando una particella incontra

la sua antiparticella si annichila, cioè sparisce, dando luogo a un raggio elettromagnetico o ad energia di altro tipo.

Un problema che rimaneva aperto dopo la scoperta della struttura dei nuclei era quello di spiegare l'origine della forza nucleare; una spiegazione soddisfacente è stata proposta nel 1935 dal fisico giapponese Yukawa sulla base dell'analogia con la forza elettromagnetica. Yukawa ha proposto che ogni protone e neutrone sia circondato da un nuovo tipo di campo con il quale esso è accoppiato attraverso una «carica» nucleare, allo stesso modo in cui un elettrone è accoppiato al suo campo elettromagnetico attraverso la sua carica elettrica. Questo nuovo tipo di campo avrà dei ◆quanti◆ di eccitazione analoghi ai fotoni e lo scambio di questi quanti fra due particelle produce la forza di attrazione. Questi quanti vennero chiamati mesoni π e vennero rivelati sperimentalmente solo nel 1946. Misure sperimentali hanno mostrato che la carica, e perciò l'interazione nucleare, è all'incirca duemila volte piú grande di quella elettrica: per questo motivo l'interazione nucleare ha preso il nome di interazione forte.

La teoria di Yukawa non spiegava però tutte le interazioni nucleari. Non spiegava ad esempio il cosiddetto decadimento β, che consiste nella trasformazione di un neutrone in un protone, un elettrone e una particella nuova, priva di carica e di massa estremamente piccola se non esattamente nulla, battezzata neutrino. Fermi spiegò il decadimento β mediante l'introduzione di una nuova forza fondamentale che regolava il processo di disintegrazione del neutrone. Questa interazione venne chiamata debole in opposizione all'interazione forte. Ne segue che, in analogia alle interazioni elettromagnetiche e forti, anche l'interazione debole deve essere mediata da un nuovo tipo di particella che venne battezzata W. Purtroppo, a differenza di quanto è successo per il mesone π di Yukawa, a tutt'oggi non si ha alcuna verifica sperimentale diretta dell'esistenza della particella W: ciò si ritiene sia dovuto al fatto che, avendo la nuova particella una massa prevista molto grande, non è possibile la sua produzione con le energie a disposizione negli attuali acceleratori. La forza debole fa capire molto bene come le forze della natura siano strettamente connesse alla struttura della materia, in quanto la forza debole non solo agisce sulla materia ma ne determina la forma, perché converte neutroni in protoni, elettroni e neutrini.

La costruzione di nuovi acceleratori ha portato alla scoperta di un numero sempre crescente di nuove particelle, sia pure tutte molto instabili, fino ad arrivare a qualche centinaio. Partiti quindi negli anni '30 da una situazione di soddisfacente semplicità con poche particelle elementari, si era arrivati dopo poco piú di vent'anni a una situazione abbastanza confusa.

Tuttavia, all'inizio degli anni '60, Gell-Mann e Ne'eman imprimono una svolta significativa nel panorama delle particelle elementari, partendo da considerazioni matematiche legate al cosiddetto gruppo $SU(3)$. Secondo la teoria da loro proposta, tutte le particelle che ubbidiscono alle interazioni forti sono composte da tre particelle «elementari» che battezzarono *quarks* (il nome è preso dal *Finnegan's Wake* di James Joyce). Molto presto però si dovette ammettere l'esistenza di un quarto quark; non solo, ma a ciascun quark si dovette in seguito associare una nuova grandezza (alla quale fu dato il nome di «colore») capace

di assumere tre «valori» o «segni» diversi (che furono indicati come rosso, verde e blu).

Il punto debole di questa teoria, che ha incontrato grande successo e popolarità, è che mai nessuno è riuscito ad osservare un quark. È possibile allora che i quark esistano davvero ma che non si riuscirà mai a vederli? Una possibilità è che, per rompere una particella ed estrarre i quark, sia necessaria un'energia talmente elevata da portare necessariamente alla creazione di una coppia quark-antiquark e successivamente alla formazione di due mesoni. In questa ipotesi non risulterà mai possibile isolare un singolo quark. La situazione è tuttavia lungi dall'essere chiara, tant'è vero che alcuni sostengono di essere riusciti a osservare un quark libero, mentre altri affermano la necessità di spiegare le proprietà delle particelle osservate senza introdurre l'ipotesi dei quark. Ad ogni modo, secondo le teorie correnti, i quark stanno assieme in quanto posseggono un nuovo tipo di «carica» e interagiscono attraverso un nuovo tipo di quanto. Questo quanto è stato chiamato gluone, mentre il nuovo tipo di carica non è altro che il colore, il quale, come s'è visto sopra, a differenza della carica elettrica è di tre tipi. In analogia con l'elettrodinamica quantistica, la dinamica a cui obbediscono queste particelle è stata chiamata cromodinamica.

Come si vede, si è ancora lontani dall'aver ridotto a pochi il numero dei costituenti elementari della materia. Partiti agl'inizi degli anni '60 con tre quark, si è già arrivati a sei, ciascuno con il rispettivo antiquark. A questi vanno aggiunte altre sei particelle che non risentono dell'interazione forte, i leptoni (elettrone, muone e leptone τ, e i corrispondenti neutrini), piú le relative sei antiparticelle.

Per quanto riguarda le forze, esse si possono far derivare da quattro interazioni fondamentali: la gravitazionale, la debole, l'elettromagnetica, la forte. Recentemente Weinberg e Salam sono riusciti a formulare una teoria in cui il campo elettromagnetico e il campo debole sono semplicemente le componenti di un'unica entità: il campo elettrodebole.

Questo successo sembra ridare fiducia al tentativo di procedere verso una descrizione unitaria di tutte le interazioni; anche se i quanti relativi alle interazioni deboli non sono ancora stati osservati, si pensa, con le nuove macchine acceleratrici in costruzione, di raggiungere le energie sufficienti per poterlo fare. Sono in corso inoltre tentativi per unificare alle precedenti anche l'interazione forte e quella gravitazionale. Questi tentativi si basano su una sempre maggiore geometrizzazione della ◆fisica◆ nello spirito della teoria del campo unificato di Einstein. Il campo che maggiormente attira l'interesse dei fisici è il cosiddetto campo di Yang-Mills; in linguaggio geometrico, esso rappresenta una connessione nello spazio fibrato principale in cui lo spazio di base è il cronotopo e il gruppo compatto svolge il ruolo di gruppo strutturale. Per il successo di queste teorie unificanti è cruciale riuscire a provare che il protone è una ◆particella◆ instabile che decade in piú particelle leggere. La sua vita media prevista è attorno ai 10^{31} anni e numerosi esperimenti sono in corso nella speranza di osservare il decadimento del protone. Un risultato positivo di questi esperimenti costituirebbe un assai serio presupposto per l'effettiva costruzione di una teoria di cam-

po unificata di tutte le interazioni, teoria che era stata l'ultimo sogno irrealizzato di Einstein e che getterebbe anche una notevole luce sulla nostra conoscenza dell'universo. Va però ricordato che le due grandi teorie della fisica moderna, la quantistica e la ♦relatività♦, che costituiscono il punto di partenza di ogni tentativo rivolto alla formulazione di una teoria unificata di tutte le forze della fisica, pur avendo portato a grandissimi successi sono ancora lontane dal potersi considerare definitive. In particolare va ricordato che in entrambe le teorie compaiono alcune caratteristiche difficoltà che sembrano direttamente legate ai fenomeni che si manifestano nell'interazione fra due particelle a distanze assai piccole e in intervalli di tempo assai brevi. La soluzione di tali difficoltà, che avrà certamente un'influenza diretta sulla comprensione della «struttura» delle particelle elementari, sembra essere connessa all'introduzione, in modo adeguato, nella teoria di un intervallo di tempo elementare (e conseguentemente di una lunghezza fondamentale), a cui è anche legata la possibilità di superare il cosiddetto problema delle divergenze dell'attuale elettrodinamica e delle altre teorie da campo. È per questa ragione che negli ultimi anni molti sforzi sono stati dedicati dai fisici alle teorie dei campi reticolari o ad altre analoghe.

I brevi cenni or ora esposti bastano a chiarire che la questione delle particelle «elementari» è ancora aperta. È possibile che la struttura gerarchica della materia – ♦atomo e molecola♦, nucleo, neutrone, quark – continui all'infinito? È stata anche fatta l'osservazione che tutta la materia potrebbe essere composta da poche particelle: i quark fondamentali, i leptoni carichi e i neutrini. Perché allora tanta sovrabbondanza di particelle? Si deve forse pensare che anziché attenersi ad un criterio di semplicità la Natura produce tutto ciò che può produrre e non solo quello di cui necessita? Spesso la situazione attuale nel campo delle particelle elementari è stata paragonata a quella che esisteva nella chimica prima della classificazione di Mendeleev del sistema periodico degli elementi. È possibile che il paragone sia appropriato e che procedendo lungo la strada attuale si arrivi a una chiarificazione soddisfacente. È però anche possibile che la strada da seguire sia completamente diversa e, per ora, imprevedibile. Altri paragoni calzanti si possono trovare nella storia della ♦fisica♦, come ad esempio il lungo cammino percorso per arrivare a spiegare il moto dei pianeti. Ogni nuova piú accurata osservazione sperimentale portava ad introdurre concetti nuovi o modifiche al complesso armamentario di cicli, equanti, ecc. usato per descrivere il moto dei pianeti. La soluzione doveva essere, come si è visto poi, di natura completamente diversa.

In effetti lo studio delle proprietà della materia a partire dalle proprietà e dal comportamento dei suoi costituenti elementari può presentare delle difficoltà scoraggianti, anche se risultati notevoli sono stati ottenuti nella comprensione delle proprietà globali dei solidi, dei liquidi, dei gas e delle varie specie di ♦plasma♦, vale a dire di quelli che vengono chiamati i quattro stati della materia inanimata. Come è possibile ad esempio prevedere quegli straordinari fenomeni come la crescita e l'autoriproduzione di un organismo biologico a partire dal comportamento delle ♦molecole♦ che lo compongono? Queste difficoltà traggono origine, in parte rilevante, anche dal fatto che le equazioni matematiche, in cui la

fisica si esprime, sono di una complessità tale che ben difficilmente si riesce a trovarne una soluzione. Per avere un'idea di queste difficoltà basti pensare che dopo trecento anni che Newton ha scritto le equazioni fondamentali della dinamica, ancora oggi ben poco si sa sul carattere stesso delle loro soluzioni. Nemmeno il problema di sole tre particelle in interazione ha trovato soluzione in generale e, nonostante gli sforzi di astronomi e matematici, non si ha nemmeno la prova della stabilità del sistema solare.

Uno dei grandi sforzi di fisici e matematici in questo secolo ha ruotato intorno alla coppia ♦reversibilità/irreversibilità♦, cercando di conciliare il determinismo e la reversibilità delle equazioni di Newton con il comportamento statistico ed irreversibile dei sistemi fisici. Inizialmente questo obiettivo si riteneva legato al passaggio dalle leggi della dinamica di un sistema a molti gradi di libertà a quelle della termodinamica attraverso una definizione statistica dell'♦entropia♦. Notevoli progressi sono stati fatti in questo campo negli ultimi vent'anni, che hanno riproposto il problema in termini completamente nuovi e su basi matematiche rigorose. In particolare è stata scoperta l'esistenza di un nuovo tipo di ♦moto♦, il cosiddetto moto stocastico (intrinseco); con questo termine viene indicato quel particolare moto di un sistema dinamico che, pur essendo governato da leggi rigorosamente deterministiche, si svolge come se il sistema fosse sotto l'azione di forze puramente casuali. Il moto stocastico è perciò altamente caotico, irregolare ed imprevedibile. È attraverso esso che è possibile gettare un ponte tra le leggi statistiche della fisica e le leggi dinamiche che un tempo erano considerate fra loro contraddittorie; non solo, ma tale moto fornisce la possibilità di derivare le prime dalle seconde. Una caratteristica tipica del moto stocastico è l'estrema instabilità delle soluzioni al variare della condizione iniziale o sotto l'effetto di perturbazioni esterne. Basti pensare che il moto di una molecola in un gas in condizioni normali di temperatura e di pressione risulta completamente alterato dopo che essa ha subito meno di cinquanta collisioni se si tiene conto dell'effetto perturbativo prodotto dal campo gravitazionale di un elettrone posto all'altra estremità dell'universo, e cioè a una distanza di circa 10^{10} anni luce.

Queste brevi considerazioni sono sufficienti ad illuminare e anche a ridimensionare l'importanza relativa del determinismo e della reversibilità delle leggi della fisica classica: basta il piú piccolo errore nel dato iniziale per avere soluzioni completamente diverse. Con una grande visione anticipatrice Maxwell affermava piú di un secolo fa: «È una dottrina metafisica che dagli stessi antecedenti seguono le stesse conseguenze. Nessuno può dubitare di questo. Ma ciò non è di molta utilità in un mondo come il nostro dove gli *stessi* antecedenti non si verificano mai e niente capita due volte». Egli invitava perciò i cultori delle scienze fisiche a studiare le singolarità e le instabilità delle cose affinché «la promozione della conoscenza naturale possa tendere a rimuovere il pregiudizio in favore del determinismo che sembra derivare dall'assumere che la scienza fisica del futuro non sia che un'immagine ingrandita di quella del passato».

Ora, le barriere psicologiche, ancora esistenti e derivanti da una tradizione di secoli che ha considerato come in opposizione i due concetti di determinismo e casualità, stanno finalmente cedendo. La moderna teoria ergodica e la teoria

della complessità algoritmica iniziata da Kolmogorov sono in grado di mostrare che il moto di certi sistemi dinamici, completamente deterministici, è indistinguibile da un moto casuale, qualunque sia il significato che si voglia dare a tale termine. In termini della dinamica simbolica una tale proprietà è espressa dal fatto che l'insieme delle orbite del sistema dinamico è *completo*. Per un esempio molto semplice ma istruttivo si consideri la trasformazione dell'intervallo [0, 1] in se stesso

$$(1) \qquad x_{n+1} = 2x_n \qquad (\text{mod } 1)$$

dove (mod 1) indica la sottrazione della parte intera in modo che x_{n+1} sia sempre compreso nell'intervallo [0, 1]. L'equazione (1) è deterministica e la sua soluzione è $x_n = 2^n x_0$ dove x_0 è il punto iniziale compreso in [0, 1]. Si noti che scrivendo x_0 in notazione binaria

$$(2) \qquad x_0 = 0,\ a_1 a_2 a_3 a_4 \ldots \qquad (\text{tutti gli } a_i \text{ sono o zero o uno})$$

la soluzione della (1), che ha come condizione iniziale x_0, si ottiene semplicemente spostando, nella (2), la virgola di un passo a destra per ogni iterazione. Poiché le successioni (2) sono in corrispondenza biunivoca con tutte le possibili successioni ottenute con il lancio della moneta (zero ≡ testa, uno ≡ croce), la soluzione della (1) può essere vista come il risultato di una successione di lanci della moneta. Piú precisamente, gli sviluppi della teoria della complessità algoritmica hanno mostrato che quasi tutte le successioni come la (2) hanno complessità di Kolmogorov positiva e sono puramente casuali in quanto soddisfano al test universale di casualità. Perciò le soluzioni della (1) sono genuinamente *random* pur essendo la (1) un'equazione rigorosamente deterministica.

Il fatto che una determinata successione abbia complessità positiva implica che l'informazione contenuta nella successione stessa non può essere compressa: cioè l'unico modo per dare la successione è quello di scriverla per intero. Ne segue che l'informazione contenuta in quasi tutte le successioni (2) è infinita e di conseguenza quasi tutti i numeri non possono essere scritti da nessun essere umano; sono esclusi ovviamente tutti i numeri con complessità di Kolmogorov nulla come gli interi, i razionali e gli irrazionali «semplici» come π ed e (base dei logaritmi naturali). La storica opposizione tra discreto e continuo riappare in maniera prepotente.

È noto come la continuità dei numeri reali sia la causa di molti problemi della ◆fisica◆. Le osservazioni sopra fatte hanno già portato alcuni (Ford, Rössler) ad ipotizzare l'esistenza di un limite superiore naturale alla precisione di una osservazione, che, unito alle proprietà stocastiche o instabili del moto, porterebbe all'assenza di determinismo nel mondo fisico. Questa linea d'indagine potrebbe portare a notevoli progressi, cosí come è sempre successo nella storia della fisica tutte le volte che si sono imposte delle limitazioni ai valori che possono assumere certe grandezze fisiche (energia, velocità della luce, ecc.).

Per concludere, si osservi che se nei due millenni precedenti sono stati compiuti enormi progressi nel campo tecnico-scientifico, non altrettanto è stato fatto per le questioni piú fondamentali. Si è penetrati con l'indagine nell'interno

del nucleo e sono stati ampliati i confini dell'universo, ma il senso di sbigottimento che l'uomo moderno prova di fronte alle innumerevoli galassie che popolano l'universo, ai quasar o ai buchi neri non è minore di quello degli antichi che osservavano le misteriose comete. L'aspetto positivo è che, come dice Espagnat, per compiere progressi significativi non è necessario conoscere la «natura ultima delle cose». Per la qual cosa ben si addice il paragone di Newton che s'immaginava come un bambino intento a raccogliere sulla spiaggia ciottoli variamente colorati mentre di fronte a lui sta, inesplorato, l'immenso oceano del Reale. [P. C. e G. C.].

Billingsley, P.
1965 *Ergodic Theory and Information*, Wiley, New York.

Caldirola, P.
1974 *Dalla microfisica alla macrofisica*, Mondadori, Milano 1976².

Caldirola, P., e Loinger, A.
1979 *Teoria fisica e realtà*, Liguori, Napoli.

Einstein, A., e Infeld, L.
1938 *The Evolution of Physics. The Growth of Ideas from Early Concepts to Relativity and Quanta*, Simon and Schuster, New York (trad. it. Boringhieri, Torino 1970³).

Espagnat, B. d'
1976 *Conceptual Foundations of Quantum Mechanics*, Benjamin, Reading Mass. 1976² (trad. it. Bibliopolis, Napoli 1980).

Feynman, R. Ph.
1965 *The Character of Physical Law*, British Broadcasting Corporation, London.

Franzinetti, P.
1980 *Introduzione alle particelle elementari*, Editori Riuniti, Roma.

Heisenberg, W.
1971 *Physics and Beyond*, Harper and Row, New York.

Jammer, M.
1957 *Concepts of Force. A Study in the Foundations of Dynamics*, Harvard University Press, Cambridge Mass. (trad. it. Feltrinelli, Milano 1971).

Maiani, L.
1981 (a cura di) *Le particelle elementari*, Le Scienze, Milano.

Peierls, R. E.
1955 *The Laws of Nature*, Allen and Unwin, London (trad. it. Boringhieri, Torino 1960).

Segrè, E.
1976 *Personaggi e scoperte nella fisica contemporanea*, Mondadori, Milano.

Toraldo di Francia, G.
1976 *L'indagine del mondo fisico*, Einaudi, Torino.

Atomo e molecola

Nella chimica e nella fisica contemporanee il concetto di molecola indica una oggettività dotata di un insieme di proprietà misurabili e rappresentata da una struttura spaziale costruita in base ad alcuni precisi nessi logici, corrispondenti alle proprietà misurate. Questo modo di concepire la molecola deriva direttamente dalle concezioni del XIX secolo, in cui essa era considerata sia come la combinazione in un composto autonomo e tipico di un insieme di unità materiali piú piccole, gli atomi, sia come un raggruppamento architettonico secondo figure geometriche ben precise.

La molecola può essere definita, indipendentemente dalla nozione di atomo, come il processo al limite della divisione della materia, che può esistere allo stato libero ed è dotata di alcune proprietà. Ma nella definizione di molecola come processo al limite della divisibilità della materia, o come la piú piccola parte irriducibile nei confronti delle proprietà chimico-fisiche, si perde la distinzione con il concetto di atomo. Nello sviluppo della meccanica statistica, atomo e molecola si confondono nel concetto di particella ultima di un sistema, dotata di proprietà individuali, ma non misurabili singolarmente. Le proprietà dei sistemi macroscopici risultano essere la conseguenza dell'insieme delle proprietà individuali.

Indipendentemente da quale possa essere la costruzione della definizione di molecola, essa rappresenta per il chimico il simbolo operazionale di una massa e di una disposizione materiale, grazie al quale egli crea il suo oggetto d'indagine.

Con lo sviluppo dell'indagine chimica, la molecola tende a sfuggire al limite imposto dal concetto originale di massa minima specifica. Uno stesso insieme di atomi può essere disposto nello spazio secondo strutture diverse. Nella chimica organica, dato l'elevato numero di atomi da cui può essere costituita una molecola e di conseguenza l'aumentata flessibilità con cui essi possono disporsi nello spazio, s'incontra spesso questa situazione. Le strutture risultanti, equiprobabili rispetto alle proprietà molecolari dipendenti dalla massa, possono essere distinte in base ad altre proprietà. Alcune possono essere messe in evidenza perché deviano in modo diverso un fascio di luce polarizzata, altre formano nel corso di una reazione composti con proprietà diverse, separabili chimicamente. All'insieme di questi fenomeni viene dato il nome di isomeria. In pratica, i pesi e la composizione atomica non bastano a caratterizzare delle specie pure, dovendo prendere in considerazione le caratteristiche strutturali per rendere conto delle proprietà molecolari. L'insieme dei metodi d'indagine chimico-fisica mette in evidenza la presenza di strutture molecolari complesse, varianti tra diversi stati energetici. In questo senso la molecola può essere definita come l'insieme delle proprietà derivanti dalla configurazione attribuita ad una specie chimica. Se queste due definizioni sono interscambiabili per le piccole molecole, al di là di alcune dimensioni molecolari questi criteri tendono a sfuocarsi. Variazioni discrete nella composizione non comportano differenze significative in

qualche carattere generalmente considerato come discriminante; d'altro canto, le proprietà fisiche di alcune macromolecole sono il risultato diretto di una distribuzione statistica. Di conseguenza, s'indebolisce sempre piú la nozione di specie chimica basata su un criterio di discontinuità della materia.

Con questo non bisogna pensare che il concetto di molecola abbia perso la sua autorità nella pratica di laboratorio, dove opera efficacemente come immagine per conoscere e controllare le trasformazioni dei corpi.

Negli anni seguenti la seconda guerra mondiale, alla molecola è stata attribuita una nuova proprietà, che consiste nel descriverla come unità in grado di essere «soggetto epistemologico». I biologi teorici, infatti, non esitano ad attribuire ad alcune molecole «stereospecifiche» il potere di riconoscere la struttura dei metaboliti da inserire in una catena di reazioni. Forse in questo, che alcuni non esitano a definire come un «abuso di linguaggio», bisogna vedere essenzialmente il tentativo di allargare la definizione del concetto di reattività. La problematica riguardante il concetto di molecola non deve far dimenticare il legame esistente con il concetto di atomo.

La definizione abituale che si dà dell'atomo è il termine ultimo della divisione della materia, che conserva le proprietà caratteristiche alla sua identificazione come elemento chimico. Piú sinteticamente, si può dire che l'atomo rappresenta la piú piccola particella di un elemento che può esistere o allo stato libero o in combinazione. Questa definizione ha introdotto la distinzione, ancora usata, della materia in corpi semplici, costituiti da atomi identici, e in corpi composti, derivanti cioè dalla combinazione di atomi diversi.

Considerato per molto tempo indivisibile, legato a concezioni filosofiche che rimontano all'antica Grecia, i progressi della fisica del XX secolo ci hanno abituato a pensare l'atomo come un sistema di particelle piú piccole in equilibrio, e a distinguere le proprietà della materia in proprietà elettroniche e nucleari.

La conoscenza della struttura atomica permette di dare una definizione esatta di atomo e di molecola. L'atomo è un sistema dotato di un solo nucleo, la molecola è un sistema a piú nuclei. L'atomo e la molecola si confondono nel problema piú generale della struttura della materia, analizzata in funzione dell'energia di legame dei suoi componenti e dell'energia corrispondente al processo di osservazione a cui il sistema è sottoposto. Cosí l'atomo o la molecola saranno considerati come particelle in rapporto all'energia con cui vengono esaminati. Se essa è inferiore all'energia di legame delle particelle subatomiche, o se essa è inferiore all'energia di legame degli atomi nelle molecole, gli atomi o le molecole si comporteranno come le unità elementari del processo.

1. *La riscoperta dell'atomismo.*

La nozione di atomo e di molecola si fonda essenzialmente sul presupposto dell'esistenza di un minimo fisico, in opposizione all'ipotesi di divisibilità all'infinito, come potrebbe far supporre l'apparente omogeneità dei corpi naturali. Ci si ricollega cosí agli antichi filosofi greci. Essi vedevano nell'aspetto omogeneo dei corpi una semplice apparenza, derivante dalla composizione delle particelle ultime. L'elaborazione poetica della filosofia di Democrito si trova nel *De rerum natura* di Lucrezio.

Dimenticata nel corso dei secoli, la concezione atomistica viene ripresa nel corso del XVI secolo in opposizione alla filosofia tomistico-aristotelica. Nel 1636 a Francoforte, l'olandese Sennert pubblica il trattato *Hypomnemata physica*, in cui è descritto un sistema chimico che, in qualche modo, prefigura una concezione molecolare. Nel trattato si parla anche della distinzione tra l'estensione e la mobilità dei corpi divisibili e i corpuscoli ultimi che conservano le forme essenziali proprie alle specie dei corpi. I «prima mixta» conservano la loro individualità nel corso delle reazioni chimiche.

Il merito di aver attirato l'attenzione degli scienziati sull'atomismo va essenzialmente a Gassendi. Nella *Cinquième Contre-Méditation*, apparsa nel 1643 grazie all'incoraggiamento di Mersenne, Gassendi loda tutti coloro i quali si sforzano di rappresentare la natura delle cose, la loro forza o proprietà, per mezzo dell'estensione, la forza, il movimento, la posizione e in generale attraverso un certo numero di piccole qualità assegnate ai corpuscoli o «principî» di cui è composta ogni cosa materiale. In realtà il tentativo di Gassendi consiste nella ricerca di una sintesi tra epicureismo e cristianesimo. In questo senso egli affermava che gli atomi erano la prima cosa creata, non in numero infinito, come per Democrito, ma in numero sufficiente per costituire l'universo finito. Essi sono dotati di un inalterabile movimento che li spinge in ogni direzione nel vuoto. Nel corso delle collisioni formano le molecole, particelle identificabili da alcuni attributi. Le particelle atomiche sono dotate soltanto di forma, resistenza, grandezza minima e «peso», conseguenza quest'ultimo del movimento. Nelle molecole un'altra forza diventa importante: la forza chimica. Le parti indivisibili, gli atomi, sono, secondo Gassendi, di una piccolezza infinita, infinitamente al di sotto delle possibilità di percezione dei nostri sensi, e là dove le nostre facoltà umane trovano il limite, comincia la grande forza della natura.

Koyré [1955] ha fatto notare che, da un punto di vista epistemologico, l'atomismo di Gassendi, presupponendo dati fissi e misurabili su una materia che non è influenzata dall'osservazione, eliminava dal bagaglio scientifico termini come 'potere', 'atti', 'accidenti', 'qualità'. In questo senso gli atomi di Gassendi rappresentano un passo avanti nel processo della scienza, anche se la conseguenza fu che la fantasia umana, non potendo osservare gli infinitamente piccoli, si limitò ad osservare le forme delle sue aggregazioni, stimolata in questo dall'uso, che si andava sempre piú diffondendo, del microscopio.

La forza chimica responsabile secondo Gassendi della forma delle molecole, si esprime nella forma che viene assunta come criterio di classificazione. Gli acidi vengono definiti come poliedri con spigoli taglienti, gli alcali come sostanze porose, i composti dello zolfo come corpi ramificati. Una reazione chimica è l'interazione di un corpo angoloso con un corpo poroso.

2. *La molecola come massa.*

L'attività del chimico consiste essenzialmente nello studio delle reazioni. Nella pratica corrente di laboratorio la molecola è vista come la piú piccola quantità in cui si possa far entrare per combinazione o sostituzione un numero irriducibile di atomi di un elemento qualsiasi. Nel descrivere una reazione chimica, i composti che vi partecipano sono rappresentati da una formula. L'equazione chimica che rappresenta una reazione indica i rapporti ponderali in cui le specie reagiscono fra di loro. In questo senso è l'espressione cifrata della legge di conservazione della massa di Lavoisier.

La ricerca sistematica dei rapporti in cui reagiscono le specie chimiche nasce e si sviluppa contemporaneamente alla rivoluzione industriale. Per spiegare la formazione delle molecole viene introdotto il concetto di affinità chimica, che serve ad esprimere la tendenza di un elemento chimico a reagire con un altro per dare un composto. Una reazione chimica è paragonata al processo di saturazione di una soluzione. Ogni elemento reagisce con un altro fino a saturazione della propria affinità.

I chimici svolgono essenzialmente un'attività di analisi dei composti naturali. Spesso lavorano nei laboratori che gli stati hanno costruito per cercare di sfruttare meglio i minerali estratti dalle miniere presenti o sul proprio territorio o nei paesi colonizzati. L'indagine chimica riguarda essenzialmente le proprietà dei metalli e dei loro derivati, in primo luogo gli ossidi. Infatti un metallo, scaldato nel corso della purificazione, può reagire con l'ossigeno presente nell'atmosfera. Un fatto non era ben chiaro: alcuni metalli possono formare piú tipi di composti con l'ossigeno.

Anche il grado di saturazione di una soluzione, cioè la quantità di solvente necessaria a sciogliere una certa quantità di soluto, è una proprietà variabile, dipendente dall'affinità del solvente per il soluto. Cosí anche l'affinità chimica di un elemento poteva essere considerata come una quantità variabile in modo continuo tra due valori estremi. I composti che presentavano un rapporto fisso di combinazione erano considerati delle eccezioni. Altre proprietà, come la volatilità o la tendenza verso la precipitazione e la cristallizzazione, erano responsabili di questo comportamento anomalo.

Proust, un abile analista al servizio del governo spagnolo, analizzò sistematicamente gli ossidi del rame e di altri metalli. Nel corso delle sue ricerche, aveva preso l'abitudine di esprimere i risultati delle analisi in percentuali in peso del campione analizzato. Si accorse che i rapporti di combinazione dei metalli con l'ossigeno sono costanti, in netto contrasto con l'ipotesi di variabilità delle proporzioni. Per spiegare le sue osservazioni, attribuí ai metalli la proprietà di «affinità elettiva» nei confronti dell'ossigeno. Sia l'ipotesi delle proporzioni variabili, sia quella della costanza delle proporzioni non avevano un supporto teorico che permettesse loro d'imporsi e di generalizzare il principio a tutti i fenomeni chimici conosciuti.

La chiave teorica venne presentata da Dalton [1802], che aveva il vantaggio

di aver potuto osservare il problema da un punto di vista esterno alle diatribe delle scuole. Interessato alla matematica e alle scienze naturali, formatosi alla lettura di Priestly, Newton e degli allievi di Hitchison, la sua ambizione era di dare una formulazione matematica alle osservazioni dei fenomeni naturali.

Cominciò a studiare la meteorologia, proseguendo il suo studio sui gas e i vapori in generale. Uno dei problemi principali dell'epoca era se l'atmosfera dovesse essere considerata come un composto chimico, data la sua omogeneità, o come un miscuglio omogeneo. Dalton mostrò che il problema poteva essere risolto se si considerava che quando due fluidi elastici sono mescolati, mentre le particelle del gas A, come quelle del gas B, si respingono tra loro, non si ha azione repulsiva tra le particelle di A e di B. La pressione o il peso totale su ognuna delle particelle deriva solamente da quelle della propria specie. Questa ipotesi, conosciuta nella chimica-fisica come legge delle pressioni parziali, non era accompagnata da considerazioni teoriche sulla struttura della materia e suscitò violente critiche da parte di grandi chimici come Berthollet, Davy, Thomson. La violenza delle critiche spinse Dalton a provare la sua ipotesi, generalizzandola. Si basò sull'osservazione del comportamento dei gas nell'atmosfera e sulla loro solubilità nell'acqua, assumendo che il comportamento dei singoli gas fosse direttamente proporzionale al peso delle particelle ultime di ogni gas: i pesi atomici o molecolari nel linguaggio moderno.

Si poneva cosí il problema di misurare i pesi delle particelle ultime. Partendo da una visione meccanicista e realista degli atomi e basandosi sulle osservazioni dei rapporti di combinazione, Dalton formulò un'ipotesi sulla costituzione delle molecole.

Quando gli atomi di due elementi A e B si avvicinano reagendo, il fattore critico è rappresentato dalla repulsione degli atomi simili fra di loro piú che dall'attrazione degli atomi diversi. Se assumiamo gli atomi sferici e di uguale grandezza, dodici atomi di B possono teoricamente venire in contatto con un atomo di A. In pratica, il risultato piú frequente è la combinazione di un atomo di A con uno di B. Meno probabile, ma possibile, è la combinazione di due atomi di B con uno di A. Tendendo a respingersi, i tre atomi si disporranno lungo una retta, cioè gli atomi di B saranno il piú distante possibile. Tre atomi di B intorno ad A implicano delle forze repulsive ancora maggiori, e si disporranno ai vertici di un triangolo equilatero.

Se è conosciuto un composto chimico di A e B, esso sarà rappresentato dalla formula **AB**; se due, le formule probabili saranno AB e AB_2, e cosí di seguito.

Con questa interpretazione dei fenomeni Dalton offriva al chimico la possibilità di scrivere i composti sotto forma di formule semplici e di rappresentarli con una figura geometrica. Nello stesso tempo egli dava la chiave per una possibile interpretazione della costanza delle proporzioni. Il problema che si presentava consisteva nella generalizzazione delle prove sperimentali e nella necessità di determinare la massa minima da attribuire ad ogni elemento chimico o ad ogni molecola. In sostanza, un problema metodologico, di cui la chiave veniva fornita dallo stesso Dalton. Infatti, se i corpi si combinano secondo

proporzioni ben definite, il loro peso può essere determinato in funzione del peso di un elemento preso come campione. Dalton scelse l'idrogeno, Berzelius l'ossigeno.

Il mondo dei chimici comincia a popolarsi cosí delle formule, parole-chiave destinate a contenere nella loro rappresentazione schematica molte informazioni.

Particolare utilità presentarono tre leggi derivate direttamente dalla concezione atomistica di Dalton.

Dulong e Petit, «convinti che alcune proprietà della materia si debbano presentare sotto una forma piú semplice, ... spinti a introdurre i piú sicuri risultati della teoria atomica nello studio di alcune proprietà che possono sembrare collegate intimamente con l'azione individuale delle molecole materiali» [1819, p. 395], analizzarono il calore specifico di numerose sostanze. Moltiplicando il valore misurato per il peso atomico, si otteneva una costante. L'osservazione fu generalizzata in una legge della materia, che fu di estrema utilità nel risolvere alcune controversie dei chimici riguardanti il valore di alcuni pesi atomici. Boltzmann confermò la validità dell'assunzione di Dulong e Petit in base alla teoria cinetica dei gas. Le deviazioni trovarono una spiegazione, su una base quantistica, da parte di Einstein [1907].

Mitscherlich, partendo dall'osservazione sistematica delle forme cristalline e basandosi sull'interpretazione di Dalton, affermò che un numero uguale di atomi, combinati nello stesso modo, produce la stessa forma cristallina, e che tale forma non dipende dalla natura degli atomi, ma semplicemente dal loro numero e dal modo in cui si combinano. Anche questa legge serviva a risolvere i casi di discussione dei rapporti di combinazione degli atomi, in base alla isomorfia delle forme cristalline.

Un passo decisivo nella determinazione dei pesi atomici e molecolari fu fatto quasi contemporaneamente da Avogadro e Ampère.

L'ipotesi di Avogadro è la seguente: «M. Gay-Lussac ha mostrato... che i gas si uniscono sempre in proporzioni semplici rispetto al volume e che quando il risultato dell'unione è un gas, anche il suo volume è in una relazione semplice con quello dei suoi componenti... Bisogna ammettere che una relazione semplice debba esistere tra il volume delle sostanze gassose e il numero di molecole semplici o composte contenute» [1811, p. 58]. L'ipotesi formulata è che volumi uguali di gas diversi nelle stesse condizioni di temperatura e pressione contengono un ugual numero di molecole. Tale valore, determinato piú tardi, è conosciuto tra le costanti della chimica e della fisica con il nome di numero di Avogadro.

La conseguenza immediata è che per volumi uguali di gas il rapporto tra le densità è uguale al rapporto fra le masse. Si era trovato un metodo diretto di misura del peso molecolare, una volta scelta una sostanza di riferimento.

Il concetto di peso molecolare o di peso atomico, di grammo-molecola o mole e di grammo-atomo entrano a far parte del linguaggio della chimica. Fino alla fine del XIX secolo i pesi molecolari saranno determinati da misure di densità. Le incertezze cui molto spesso le misure davano luogo fecero dubitare della validità dell'ipotesi atomica e delle leggi che si basavano su di essa. L'affinamento

delle tecniche sperimentali ha permesso di riconoscere che le deviazioni erano dovute a fenomeni diversi come la dissociazione termica o l'esistenza di corpi sotto forme diverse, dipendenti dalle condizioni sperimentali, ma non legate ad una variazione della composizione: le specie allotropiche.

Questi fenomeni, oggi compresi completamente, hanno rinforzato l'ipotesi atomistica a tal punto che lo studio delle deviazioni è importante per osservare i fenomeni di dissociazione. Ma si sono posti altri problemi. L'identità molecolare perde le sue caratteristiche nei cristalli, dove predomina una struttura tridimensionale dei motivi costituiti dal raggruppamento di atomi o di ioni; in altri casi, mentre è possibile definire un raggruppamento atomico elementare, non è possibile stabilire se esso rappresenti un minimo fisico autonomo. È spesso il caso dei polimeri, la cui lunghezza può variare localmente. In genere si è d'accordo che una molecola possa essere isolata e fisicamente oggettivata nello stato di soluzione. Alcune proprietà delle soluzioni, come pressione osmotica, temperatura di ebollizione o di congelamento, sono espresse in funzione del peso molecolare del composto presente.

3. *Le proprietà macroscopiche come media delle caratteristiche microscopiche.*

I chimici avevano colto nel modello di Dalton la possibilità di definire quantitativamente l'unità di massa partecipante ad una reazione chimica. Ma il modello atomico aveva come punto di partenza l'identità dell'azione delle particelle simili nei confronti di una grandezza macroscopica tipica dei fluidi, la pressione. Tale modello poteva essere utile per interpretare i nuovi fenomeni legati strettamente alla rivoluzione industriale e all'introduzione della macchina a vapore.

Sadi Carnot, nelle sue *Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance* [1824], accetta implicitamente il modello atomistico di Dalton. Esplicito riferimento viene fatto alla tabella delle pressioni parziali di Dalton, dove venivano correlati la temperatura e la pressione, e alla legge di Dulong e Petit, assumendo cosí il calore specifico di un atomo come l'unità minima della quantità di calore di un fluido.

Clapeyron e Clausius svilupparono le idee di Carnot. Per determinare le funzioni di stato tipiche della termodinamica, Clausius si serví di un modello macroscopico, mentre per sviluppare la teoria cinetica dei gas, soprattutto la diffusione, utilizzò un modello molecolare. Fu in grado cosí di definire un valore numerico per il numero di Avogadro.

Le teorie cinetiche dei fluidi si propongono di spiegare le proprietà macroscopiche a partire da un'analisi statistica dei movimenti delle particelle che li costituiscono. Il punto di partenza consiste nell'analisi del fenomeno di collisione tra due particelle. Ogni molecola o atomo è assimilato ad una massa puntiforme, il cui movimento è, in un dato istante, funzione della posizione e della velocità. Dato il gran numero di molecole contenuto in un volume, anche se limitato, ci si deve accontentare di una distribuzione statistica attraverso le funzioni di distribuzione.

Maxwell partí dal presupposto che la maggior parte delle proprietà di un gas avrebbero potuto essere calcolate se, invece di determinare le posizioni e le velocità di tutte le molecole, si fosse determinato il numero medio di molecole che hanno un determinato valore di tali grandezze. L'assunzione alla base è che un gas si distribuisce uniformemente nello spazio e che, di conseguenza, il numero medio di molecole è uguale in diversi punti del volume occupato.

Boltzmann cercò, sulla base di un modello atomico della materia, la relazione tra il secondo principio della termodinamica e il principio di azione minima in meccanica. Considerò che l'energia potesse essere pensata distribuita in quanti, piccoli, ma finiti, tra le molecole. Il problema si riduceva ad un'analisi di tutti i modi in cui ad ogni molecola poteva essere associato un quanto di energia o multipli interi di esso. In sostanza un problema di analisi combinatoria.

Una grande parte del lavoro di Boltzmann fu speso nella difesa dell'ipotesi atomica della materia. Tra quelli che si opposero si trovano Mach, Ostwald, Duhem, Helm. Mach riconosceva che l'ipotesi atomica era utile allo sviluppo delle scienze ma, ancora nel 1912, metteva in guardia i fisici dal considerarli come una realtà oggettiva. Ostwald, Duhem, Helm negavano la validità dell'ipotesi atomica anche dal punto di vista di ipotesi di lavoro, proponendo un principio energetico generalizzato, basandosi sull'ipotesi di una possibile generalizzazione della termodinamica.

Cosí, nell'ultima parte del XIX secolo, ci si trova in presenza di due definizioni della molecola. Per i chimici è la piú piccola quantità di una sostanza che conserva le caratteristiche chimiche del composto e che fa parte di una reazione, per i fisici è la piú piccola parte di una sostanza che si muove come un tutto. La distinzione tra atomo e molecola non è chiara.

Combinando la concezione fisica dei movimenti molecolari con le misure chimiche riguardanti le proprietà delle soluzioni, Perrin si propose di dare al concetto di molecola una realtà oggettiva. Il problema consisteva nel trovare il limite di grandezza al raggruppamento di atomi che possa verificare le leggi delle soluzioni.

Brown aveva osservato nel 1828 che alcune particelle, sospese in un liquido, erano dotate di un movimento incessante, simile al movimento delle particelle sospese nell'atmosfera, come è possibile osservare quando un raggio di sole penetra in una stanza buia. L'idea di Perrin fu di applicare la teoria cinetica dei gas al movimento browniano. Se l'ipotesi molecolare è giusta, si sarebbe potuto determinare il numero di Avogadro, cioè la quantità di molecole contenute in una grammo-molecola. Con un procedimento ingegnoso, Perrin e Svedberg determinarono il valore di questa grandezza, che risultò coincidere con le misure effettuate sui gas. Confortato da questi risultati, con l'appoggio dell'interpretazione teorica formulata da Einstein, Perrin difese davanti ai piú grandi fisici europei dell'epoca, riuniti nel 1911 a Bruxelles per il primo congresso Solvay, la realtà oggettiva della molecola.

4. *La molecola come struttura architettonica.*

La teoria di Dalton sulla costituzione delle molecole permetteva di associare alla formula di un composto una struttura geometrica.

In genere la correlazione tra composizione atomica e struttura risultava facile per i composti della chimica inorganica, che sono costituiti da un piccolo numero di atomi. Il problema diventava molto piú complicato per i composti della chimica organica. La composizione atomica di questi composti non permetteva di associare una struttura semplice a causa del numero elevato di atomi che formano una molecola. (Il piú semplice composto della chimica organica, il metano CH_4, contiene cinque atomi).

Il lavoro di analisi sui composti organici aveva fatto sviluppare un nuovo campo d'indagine: la sostituzione, lo studio cioè della proprietà dei composti organici di sostituire facilmente uno o piú atomi, in genere quelli d'idrogeno, con altri. La sostituzione non implicava però un cambiamento radicale nelle proprietà del composto. La prima osservazione di tale fenomeno fu fortuita. Nelle prime industrie a carattere artigianale era stato osservato che il cloro, utilizzato come sbiancante, era fissato in piccole quantità dalla cera nel corso del trattamento. Si cominciarono ad osservare sistematicamente le sostituzioni, accumulando in poco tempo una quantità enorme di dati.

Il fenomeno, caratteristico dei composti organici, veniva spiegato con teorie che avevano un valore limitato per un gruppo ristretto di composti. Si era creata cosí una situazione paradossale. La chimica inorganica cercava di organizzare le conoscenze in base a leggi di portata generale, mentre la chimica organica si perdeva sempre piú nella massa dei composti che venivano isolati. D'altronde, la sintesi dell'urea, effettuata nel 1828 da Wöhler, aveva fatto crollare la distinzione esistente tra chimica organica e inorganica. Le stesse tecniche di laboratorio potevano essere utilizzate per sintetizzare sia i composti organici sia quelli inorganici.

Interessato a questa situazione, il chimico francese Laurent decise di basare il lavoro di ricerca per la sua tesi di dottorato sull'analisi della letteratura riguardante la chimica organica. Aveva ricevuto la sua formazione presso l'Ecole des Mines di Parigi, una delle piú prestigiose scuole di chimica dell'epoca. Si era familiarizzato con tutte le tecniche di analisi dei composti inorganici e con le leggi che permettevano di risalire dalla composizione atomica alla struttura. In aperto contrasto con il mondo accademico di Parigi, cui rimproverava l'autoritarismo e la chiusura verso le novità scientifiche, militante nella sinistra repubblicana, decise di lavorare da solo all'elaborazione della tesi e aprí una scuola privata di chimica a Parigi.

Nel corso del suo lavoro di ricerca si accorse che le proprietà dei composti organici variano in modo continuo in base al numero di atomi di carbonio presenti nella molecola. Assunse di conseguenza che il nucleo fondamentale dei composti organici fosse rappresentato dall'insieme degli atomi di carbonio, che disponeva secondo i vertici di una piramide. A tale struttura attribuiva una sta-

bilità tale da potersi mantenere intatta nel corso di una reazione. I composti organici povano essere classificati in base alla presenza nella loro molecola di una stessa struttura piramidale. Laurent chiamò queste strutture «radicali».

In base a questa teoria, la sostituzione trovava una interpretazione semplice e valida per tutti i composti. Se si sostituisce un atomo diverso dal carbonio, la struttura fondamentale resta immutata, quindi le proprietà del composto non variano; la sostituzione di un atomo di carbonio provoca un cambiamento radicale nella struttura, quindi un composto completamente diverso da quello di partenza.

La semplicità della teoria di Laurent facilitò la sua accettazione e diffusione tra i chimici dell'epoca. Cosí, la struttura diventa importante per interpretare le proprietà dei composti ed entra a far parte delle ricerche del chimico.

Per poter stabilire la sequenza spaziale degli atomi in un composto, bisognava definire esattamente la capacità di un elemento di combinarsi con un altro, la valenza, e accettare il principio che due atomi uguali potessero legarsi fra di loro. Il merito di aver definito questi due concetti viene attribuito a Kekule che cosí si espresse: «Si è visto che la quantità di carbonio che i chimici hanno riconosciuto come la piú piccola possibile, cioè come un atomo, si unisce sempre con quattro atomi di un elemento monoatomico o con due atomi di un elemento diatomico, in generale la somma delle unità chimiche unite con un atomo di carbonio è quattro. Questo ci porta a considerare il carbonio come tetrabasico o tetratomico. Nel caso di sostanze che contengono piú atomi di carbonio, bisogna assumere che almeno alcuni degli atomi sono legati nello stesso modo nel composto dall'affinità del carbonio e che gli atomi di carbonio si legano uno all'altro, in modo che una parte dell'affinità di uno è naturalmente ingaggiata con una parte uguale dell'affinità dell'altro. Il modo piú semplice, quindi piú probabile, di una tale associazione degli atomi di carbonio è che una affinità di uno sia legata ad una unità di un altro» [1859, p. 129].

Kekule era ossessionato dalla possibilità di legare gli atomi fra di loro in una struttura complessa. Nel 1890, durante la celebrazione in suo onore per il venticinquesimo anniversario della scoperta della formula del benzene, davanti ai membri dell'associazione dei chimici tedeschi, Kekule rivelò che l'idea di legare gli atomi fra di loro gli apparve mentre viaggiava su un omnibus a Londra, una notte d'estate del 1854 tra Islington e Clapham. «Caddi in un sogno. Gli atomi ballavano davanti ai miei occhi. Le avevo sempre viste in movimento quelle piccole cose, ma non ero mai stato capace di discernere la natura del loro movimento. Ora, tuttavia, seppi come due piccoli atomi si uniscono per formare un paio; uno piú grande poteva legarsi a due piú piccoli, come uno piú grande ancora poteva legarsi con tre o quattro atomi piú piccoli. Tutto questo mentre l'insieme continuava a danzare follemente. Seppi come i piú grandi formano delle catene e i piú piccoli si aggrappano soltanto alle estremità delle catene» [cit. in Partington 1964, p. 537]. Svegliato dal conducente dell'omnibus, passò tutta la notte a mettere nero su bianco i suoi sogni.

L'intuizione di Kekule, che aveva ricevuto una formazione di architetto prima di dedicarsi agli studi chimici, fu fondamentale nello sviluppo delle sintesi

organiche, ma le formule che proponeva non rendevano conto di un altro fenomeno: l'isomeria. Esso era stato già messo in evidenza da Faraday, Wöhler e Berzelius e, come amava dire Pasteur, il suo interesse consisteva essenzialmente nella necessità di accettare il principio che i corpi possono essere differenti per il solo fatto che la disposizione degli atomi non è la stessa in due molecole aventi la stessa composizione.

Pasteur aveva mostrato durante le sue ricerche sui composti dell'acido tartarico che esiste una relazione tra forma cristallina e attività ottica di una soluzione, facendo dipendere dalla dissimmetria molecolare dei cristalli il diverso comportamento delle molecole rispetto alla luce polarizzata.

Rifacendosi all'idea di Kekule sulla struttura dei composti organici, Le Bel e van't Hoff, indipendentemente, proposero la teoria del carbonio asimmetrico, partendo dall'ipotesi che i legami del carbonio puntano verso i vertici di un tetraedro. «Nel caso che le quattro affinità di un atomo di carbonio siano saturate da quattro gruppi differenti e univalenti, due e solamente due tetraedri differenti possono essere ottenuti, di cui uno è l'immagine speculare dell'altro..., due formule strutturali di isomeri nello spazio... Ogni composto del carbonio, che in soluzione impone una deviazione al piano di polarizzazione, contiene un atomo di carbonio asimmetrico» [van't Hoff 1875].

La costruzione grafica delle proprietà chimiche ha dato allo sviluppo della chimica organica un impulso decisivo. La struttura di una molecola viene rappresentata in modo alquanto artificioso secondo una figura geometrica. Queste formule costituiscono essenzialmente un potente aiuto tassonomico. Inoltre le proprietà che possono essere collegate alla struttura aiutano il chimico nel prendere le decisioni. In questo senso possono essere considerati degli strumenti economici di simulazione per mezzo dei quali il chimico può rappresentare le trasformazioni molecolari che disegna prima di passare alla prova sperimentale.

Kekule volle dimostrare come le formule di struttura che proponeva permettessero di predire le proprietà e la possibilità di sintesi dei composti organici. Scelse, tra varie possibilità, di sintetizzare il trifenilmetano, composto base per ottenere un colorante, il rosa anilina, e la scelta si rivelò fortunata. La ricerca sintetica nel campo dei coloranti, primo esempio di produzione industriale nel campo della chimica organica, ricevette uno stimolo considerevole. L'espansione e la superiorità acquistata dall'industria tedesca dei coloranti a partire dalla fine del XIX secolo sono dovute in parte ai chimici formatisi alla scuola di Kekule, in onore del quale l'associazione degli industriali chimici eresse una statua.

Ma da molto tempo ormai i vertici dei tetraedri non rappresentano piú crudamente la posizione degli atomi intorno al carbonio centrale, come gli atomi non rappresentano piú la piú piccola parte indivisibile della materia. L'unificazione della problematica atomo/molecola nel problema piú generale della struttura della materia è il risultato della ricerca dei fisici e dei chimici all'inizio del XX secolo. Alcuni fatti sperimentali hanno imposto il cambiamento nel modello interpretativo.

5. *L'evidenza sperimentale della complessità della struttura atomica.*

Nel 1858 Geissler, soffiatore di vetro e tecnico meccanico a Bonn, realizzò un tubo di vetro in cui, dopo aver fatto parzialmente il vuoto, faceva avvenire una scarica elettrica. Si aveva cosí la comparsa di una radiazione luminosa.

Il nome di raggi catodici deriva dalla dimostrazione di Perrin che la radiazione è costituita da corpuscoli carichi negativamente, proiettati a partire dal catodo. Perrin dichiarò che l'idea gli era venuta dallo studio dei lavori di Helmholtz sull'elettrolisi. Volendo interpretare il nuovo fenomeno secondo la teoria corpuscolare, si accorse che era necessario introdurre il concetto di atomo di elettricità. In questo non bisogna vedere niente di rivoluzionario, ma semplicemente il tentativo d'interpretare secondo la teoria atomista quello che Ostwald e altri ricercatori interpretavano secondo la teoria energetica. Infatti, il concetto di minimo di elettricità è contenuto implicitamente nelle leggi di Faraday sull'elettrolisi. Faraday aveva dimostrato che il rapporto tra le quantità di sostanze che si depositano agli elettrodi al passaggio di una stessa quantità di elettricità è uguale al rapporto in cui le stesse sostanze si combinano con l'idrogeno.

Restavano da determinare le caratteristiche delle particelle costituenti i raggi catodici. Il primo ad accorgersi che la massa di tali particelle è molto piccola fu Wiechert, seguito da Thomson. Quest'ultimo misurò nel 1898 il rapporto carica/massa. Il valore della carica fu determinato in modo definitivo tredici anni dopo da Millikan, con un metodo abbastanza ingegnoso. Egli, infatti, ionizzava una goccia d'olio sospesa tra le due armature di un condensatore, cioè osservava il movimento della goccia ionizzata in un campo elettrico. L'osservazione continuava per molte ore con l'aiuto di un microscopio. Dal movimento osservato, e conoscendo il campo elettrico, è possibile risalire alla carica.

Thomson osservò che le particelle prodotte in tali tubi erano le stesse indipendentemente dal gas presente e dalla natura degli elettrodi. Queste particelle, inoltre, avevano le stesse caratteristiche di quelle emesse dai metalli quando sono scaldati (effetto Edison) o in seguito all'azione dei raggi ultravioletti e X sui metalli. La conclusione logica fu che queste particelle, gli elettroni, sono i costituenti comuni a tutti gli atomi.

Nel 1895 il fisico tedesco Roentgen osservò accidentalmente che una lastra ricoperta di platino-cianuro di bario diventa luminosa nelle vicinanze di un tubo a raggi catodici contenuto in una scatola di cartone nero in una stanza al buio. Interessato dal fenomeno, notò che le lastre fotografiche al riparo dalla luce venivano impressionate e che poteva essere fotografata una mano mettendo in evidenza le ossa. Concluse che il tubo emetteva una specie di radiazione molto potente e, non sapendo come chiamarla, la chiamò X.

Nel 1896 Perrin, basandosi su un'ipotesi di Poincaré, identificò la sorgente dei raggi X nell'ostacolo che gli elettroni incontrano sul loro cammino. L'intensità della radiazione emessa dipendeva dalla natura dell'ostacolo, permettendo cosí di stabilire una relazione tra peso atomico e intensità dei raggi X.

Ma i raggi X avevano un comportamento strano. Non erano deviati né da un campo elettrico né da un campo magnetico, non erano né riflessi né rifratti, non presentavano fenomeni di diffrazione come qualsiasi sorgente luminosa quando i raggi sono fatti passare attraverso una fessura. Essi venivano solamente assorbiti, di piú o di meno.

Se si attribuiva ai raggi X una natura ondulatoria, come la luce, la lunghezza d'onda doveva essere necessariamente molto piccola. Per verificare tale ipotesi bisognava eseguire una misura di diffrazione, ma per poter far questo si doveva trovare un reticolo che contasse qualche migliaio di righe per millimetro, mentre il miglior reticolo che si era capaci di costruire ne conteneva 500 per millimetro.

Max von Laue pensò che un cristallo doveva costituire il reticolo ideale per osservare la diffrazione dei raggi X. Il cristallografo francese Häuy, basandosi sull'osservazione che i frammenti di un cristallo conservano la stessa forma del cristallo iniziale, aveva formulato l'ipotesi che la struttura geometrica dei cristalli fosse dovuta alla ripetizione nello spazio di una struttura fondamentale, il nucleo cristallino.

Bravais si basò su questa ipotesi e classificò i minerali secondo poliedri, la cui ripetizione nello spazio dà origine alla forma cristallina osservata. Se i vertici di tali poliedri vengono fatti coincidere con la posizione delle molecole del composto, un cristallo può essere rappresentato come una rete tridimensionale, le cui maglie sono riconducibili a figure geometriche semplici. Se il cristallo ha la forma di un cubo, il reticolo può essere visto anche come un insieme di linee parallele.

Le dimensioni di una molecola sono molto piccole, e la distanza intermolecolare in un reticolo deve essere piccola. La sequenza delle molecole in un cristallo rassomiglia cosí ad un reticolo di diffrazione con un numero di righe che non si sarebbe mai stati capaci di realizzare in laboratorio.

L'esperienza fu eseguita da due collaboratori di Laue, Friedrich e Knipping, confermando pienamente l'ipotesi. Il metodo fu perfezionato dai Bragg e da De Broglie, che stabilirono una relazione matematica tra la distanza degli atomi nei reticoli, la lunghezza d'onda dei raggi X e l'angolo d'incidenza. Debye e Scherz estesero il metodo ai sistemi amorfi e Davisson e Germer, per caso, osservarono che anche un fascio di elettroni, comportandosi come un'onda luminosa, veniva diffratto dalle sostanze che incontrava, anche se queste si trovavano allo stato gassoso.

La struttura di una molecola è pensata come un ordinamento tridimensionale di atomi. Se non è dotata di struttura cristallina, può essere assimilata ad un composto amorfo. La distanza tra gli atomi costituenti la molecola serve come reticolo base di diffrazione. Facendo incidere un fascio di raggi X di lunghezza d'onda conosciuta, si possono ricavare le distanze inter-reticolari, cioè le distanze tra gli atomi. Si era trovato cosí il metodo per «fotografare» la molecola.

A partire da questo momento le strutture tridimensionali delle molecole non sono soltanto un'ipotesi di lavoro ma una realtà. L'affinamento dei metodi sperimentali permette di risolvere delle strutture complesse, spesso in modo

completamente automatico. I raggi diffratti vengono automaticamente trasformati in informazione per un elaboratore elettronico, cioè al posto della fotografia si hanno delle schede perforate. Programmi di calcolo standard risolvono la struttura. Nel 1952 Watson e Crick riuscirono a dedurre, a partire da esperienze di diffrazione, la struttura di una macromolecola complessa, il DNA. La biologia molecolare trovava l'oggettivazione del suo soggetto di lavoro.

In un tubo a raggi catodici, se si pratica un foro sul catodo, si osserva il passaggio di una radiazione chiamata da Goldstein, che la osservò per primo, 'raggi canale'. Nel 1901 Wien osservò sperimentalmente che essi erano costituiti da particelle con una carica positiva e che la loro massa era di molto superiore alla massa dell'elettrone e vicina alla massa degli atomi del gas presenti nel tubo.

Thomson notò che i raggi canale, deviati da un campo elettrico o magnetico, disegnavano su uno schermo luminoso, perpendicolare alla loro direzione iniziale, una parabola dai contorni talmente netti, che si poteva assegnare una parabola per ogni elemento chimico. Nel tentativo di fare questa assegnazione, Aston, lavorando nello stesso laboratorio di Thomson, perfezionò l'apparecchio in modo che i raggi deviati si concentrassero in un punto. Era nata cosí la spettroscopia di massa, ampiamente utilizzata nella fisica delle particelle.

Ma quale fu la sorpresa di Aston nel constatare che campioni omogenei di elementi chimici davano luogo a piú fasci canale deviati. Avendo la stessa carica, l'unica spiegazione possibile consisteva nella differenza di massa. Si ritrovava un risultato noto ai fisici che s'interessavano di radioattività: l'esistenza di isotopi.

Crollava cosí uno dei principî fondamentali dell'atomismo, che considerava uguali tutti gli atomi di uno stesso elemento. Nello stesso tempo, però, si risolvevano tutte le discrepanze nella determinazione dei pesi molecolari e dei calori specifici atomici.

Interessato ai nuovi fatti della fisica, Poincaré suggerí a Becquerel di studiare i raggi X utilizzando composti diversi. Il caso fece scegliere a Becquerel un sale d'uranio: con sua grande sorpresa osservò che il cristallo non solo emetteva una radiazione capace d'impressionare una lastra fotografica, ma che rendeva un gas conduttore di elettricità.

Becquerel affidò ai coniugi Curie il compito di proseguire lo studio e di isolare altri composti che fossero dotati delle stesse proprietà. Nel giugno del 1898 i due ricercatori isolarono un nuovo elemento, il radio, in dicembre dello stesso anno il polonio.

Rutherford analizzò in dettaglio le radiazioni emesse e notò che soltanto due di esse erano deviate da un campo magnetico. Chiamò i tre fasci in cui si divideva la radiazione raggi α, β, γ. Le particelle α furono assimilate a nuclei di elio carichi positivamente e l'osservazione sistematica permise di stabilire che l'emissione di una particella α provoca l'abbassamento del numero atomico di due unità, mentre l'emissione di raggi β conserva il peso atomico, ma aumenta il numero atomico di una unità.

Si poneva cosí ai fisici il problema di trovare un modello dell'atomo che permettesse di descrivere i nuovi fatti osservati.

Cosí si espresse Thomson: «Considero l'atomo come contenente un ampio numero di corpuscoli... Nell'atomo normale questo insieme forma un sistema che è elettricamente neutro. Benché i corpuscoli isolati si comportino come ioni negativi, quando essi sono associati nell'atomo neutro l'effetto negativo è bilanciato da qualcosa che fa sí che lo spazio, nel quale i corpuscoli si trovano dispersi, si comporti come se fosse dotato di una carica elettrica positiva uguale come quantità alla somma delle cariche negative dei corpuscoli» [1899, p. 547].

L'atomo, secondo Thomson, consiste in una sfera carica positivamente, di raggio uguale al raggio atomico (10^{-8} cm). Gli elettroni all'interno di questa sfera eseguono delle oscillazioni armoniche ed emettono una radiazione di frequenza data. Lorentz, prendendo spunto dai lavori di Maxwell, formulò il modello matematico basato sui principî dell'elettrodinamica classica.

Rutherford pensò di poter misurare la carica positiva presente nell'atomo in base alle deviazioni di una particella α nelle vicinanze del nucleo, secondo la legge di repulsione di due cariche positive formulata da Coulomb. Furono cosí sistematicamente osservate le deviazioni delle particelle α nelle vicinanze dei nuclei di varie sostanze. I risultati sperimentali potevano essere interpretati solamente se si pensava la carica positiva come concentrata in un volume piú piccolo del volume atomico.

Rifacendosi ad un modello proposto nel 1903 da Nagoaka, Rutherford propose per l'atomo un modello planetario: il nucleo al centro, gli elettroni ruotanti su orbite circolari.

6. *La quantificazione della radiazione può spiegare la stabilità dell'atomo.*

Benché riuscisse a spiegare molti fenomeni sperimentali, l'atomo di Rutherford non era dotato di una struttura stabile. L'elettrone, secondo la teoria elettromagnetica classica, muovendosi intorno al nucleo emette energia, diminuendo cosí progressivamente il raggio della sua orbita fino a cadere sul nucleo. La soluzione fu trovata nell'ipotesi di una struttura discontinua della radiazione.

L'idea di quantizzare la radiazione era stata introdotta nella fisica agli inizi del XX secolo: nel 1900 Planck, per interpretare la radiazione del corpo nero, era dovuto ricorrere all'ipotesi che l'emissione della radiazione avvenisse solo per quantità discrete. I motivi che lo avevano spinto ad introdurre quest'ipotesi sono spiegati dallo stesso Planck in una lettera a Wood: «Boltzmann ha spiegato come un equilibrio termodinamico si stabilisca per mezzo di un equilibrio statistico e, se un tale tipo di approccio è applicato all'equilibrio tra materia e radiazione, si trova che la perdita continua di energia sotto forma di radiazione può essere eliminata assumendo che l'energia, al momento dell'emissione, sia obbligata a restare sotto forma di quanti. Questa era semplicemente un'assunzione formale e in realtà non ho voluto darle molta importanza eccetto per il fatto che, costi quello che costi, dovevo trovare un risultato positivo» [citata in Hermann 1971].

Questo «atto di disperazione», come lo definiva lo stesso Planck, gli era costato la rinunzia all'opposizione praticata per molto tempo contro l'atomismo e alla polemica contro Boltzmann.

Nel volume XVII degli «Annalen der Physik», Einstein [1905], ancora praticamente sconosciuto, pubblica un articolo dal titolo *Su un punto di vista euristico concernente la produzione e la trasformazione della luce*. Se Planck aveva supposto che lo scambio di energia tra materia e radiazione avvenisse per quantità discrete, Einstein attribuisce alla radiazione una struttura discretà, i fotoni, cui è associata un'energia uguale a $h\nu$. La costante h è la stessa introdotta da Planck per spiegare la radiazione del corpo nero ed è diventata in seguito una costante universale con il nome di costante di Planck.

Un giovane fisico viennese, Haas, cercò di trovare un significato fisico alla costante di Planck. Pose cosí l'uguaglianza tra l'energia dell'elettrone nell'atomo d'idrogeno e la costante di Planck. Riusciva a trovare alcune grandezze atomiche, ma sfortunatamente il modello su cui si basava era quello di Thomson.

Il problema dei quanti passò da problema marginale – in pratica le discussioni sui quanti coinvolgevano soltanto una parte dei fisici dell'epoca e soprattutto i fisici tedeschi – a problema centrale della fisica con il congresso Solvay del 1911. Durante questo congresso vennero discussi la realtà oggettiva della molecola, l'idea dei quanti, i primi modelli atomici. In pratica furono analizzati i nuovi problemi della fisica teorica.

Un giovane fisico danese, Niels Bohr, lesse attentamente gli atti del congresso ed ebbe la possibilità di discuterne frequentemente con uno dei protagonisti, Rutherford, presso il quale passava l'anno di ricerche all'estero, obbligatorio per chi avesse sostenuto una tesi di dottorato. Nella preparazione della tesi, centrata sullo studio delle proprietà elettroniche dei metalli, Bohr si era convinto che l'elettrodinamica classica era inadeguata a spiegare i nuovi fenomeni atomici. Leggendo gli atti del congresso venne in contatto con il modello di Haas e pensò di applicare la quantificazione al movimento degli elettroni nell'atomo. Ogni movimento avrebbe potuto essere caratterizzato da un'energia direttamente proporzionale a multipli interi di h. Combinando tale principio con le espressioni classiche del moto, si sarebbero potute stimare, con buona approssimazione, le forze agenti in un sistema atomico.

Applicando le sue ipotesi al modello di Rutherford in maniera qualitativa, Bohr si accorse che molte cose diventavano chiare. Gli atomi potevano essere pensati come un nucleo contornato da anelli su cui si muovono gli elettroni, ma anche le molecole potevano essere pensate come un sistema a piú nuclei, con gli elettroni ruotanti in anelli simmetrici rispetto alla disposizione dei nuclei. Si introduceva cosí, per la prima volta, una distinzione netta tra atomo e molecola. Il legame chimico poteva essere considerato come il risultato della messa in comune degli elettroni fra gli atomi, come alcuni calcoli semplici facevano prevedere per la molecola d'idrogeno.

Ritornato a Copenhagen, Bohr si propose di pubblicare questi risultati cercando una base piú solida ai suoi principî. Discutendo con uno dei suoi colleghi, lo spettroscopista Hansen, Bohr apprese che gli spettroscopisti usavano

per indicare la regolarità degli spettri degli elementi delle formule empiriche molto semplici, generalmente funzione della differenza dell'inverso dei quadrati di due numeri interi moltiplicato una costante. L'espressione matematica era molto simile a quella usata da Bohr, che pensò di servirsene per la sua teoria.

Nel 1913 Bohr presentò la sua teoria atomica, basandosi su due postulati. Il primo assume che gli elettroni si trovino su stati stazionari, le orbite, il cui comportamento può essere descritto tramite le equazioni classiche del movimento; il secondo stabilisce che le transizioni tra due stati stazionari non sono processi descrivibili classicamente e che sono accompagnate dalla emissione o dall'assorbimento di un quanto di radiazione omogenea o da multipli interi [Bohr 1913].

7. *Il modello atomico permette di spiegare la formazione dei legami chimici.*

Il modello di Bohr diventava lo schema teorico che permetteva di analizzare la struttura atomica e i fenomeni spettrali. Nello stesso articolo, nella terza parte esattamente, Bohr applicò i principî del suo modello per cercare di spiegare la regolarità del sistema periodico degli elementi e la formazione delle molecole.

Da alcuni anni infatti i fisici avevano cercato di stabilire un legame tra struttura dell'atomo e due fenomeni chimici fondamentali: il legame chimico, cioè la capacità degli atomi di legarsi fra di loro, e la valenza, la proprietà che determina il numero di atomi o gruppi atomici con cui un atomo isolato o un gruppo di atomi si unisce chimicamente, oppure la capacità di un elemento di combinarsi con l'idrogeno.

Thomson, nel suo trattato *Corpuscular Theory of Matter* [1907], aveva postulato che il legame chimico fosse basato essenzialmente sul trasferimento degli elettroni da un atomo ad un altro, con conseguente attrazione elettrostatica.

Secondo Kossel, allievo di Sommerfeld, in accordo con alcune idee espresse da Bohr, le proprietà chimiche di un atomo sono determinate dal numero di elettroni nell'anello piú esterno, gli elettroni di valenza. Egli postulò, inoltre, che la notevole stabilità, dal punto di vista della reattività chimica, dei gas nobili fosse dovuta al completo riempimento dell'orbita esterna da parte degli elettroni. Questi erano otto per tutti i gas nobili eccetto l'elio, che ne aveva solamente due. Ogni elemento, per raggiungere una configurazione elettronica stabile, corrispondente alla configurazione elettronica del gas nobile che lo precede o lo segue nella tavola periodica degli elementi, perde o acquista elettroni. Principale responsabile di un legame chimico è l'attrazione elettrostatica, anche se non si può escludere una forma di compartecipazione degli elettroni di due atomi, soprattutto quando essi sono molto simili.

Un mese dopo la pubblicazione del lavoro di Kossel, il chimico americano Lewis pubblicò un lavoro in cui, volendo spiegare i legami presenti nei composti organici, affermava che il legame piú comune della chimica era la conseguenza della messa in comune di due elettroni tra due atomi. Per spiegare

la possibilità degli atomi a formare i legami, Lewis immaginava gli elettroni situati ai vertici di un cubo, con il nucleo al centro.

Le idee di Kossel e di Lewis rimasero sconosciute ai chimici. La spiegazione comunemente accettata è che i chimici, impegnati nei problemi di difesa a causa della guerra, furono sviati dai problemi teorici.

Nel 1919 Langmuir, un chimico americano, propose una sintesi tra le teorie di Lewis e Kossel. I legami chimici possono essere di due tipi: 1) elettropolare, conseguenza del trasferimento di un elettrone da un atomo ad un altro e dotati di energia essenzialmente uguale all'attrazione elettrostatica tra i due ioni; 2) non-polare, dovuto alla messa in comune degli elettroni fra i due atomi.

Gli atomi possono essere rappresentati dal loro simbolo circondato da tanti puntini quanti sono gli elettroni presenti. Le formule di struttura molecolari sono pensate in modo che ogni atomo, eccetto l'idrogeno, si trovi ad essere circondato da otto elettroni.

La teoria riusciva a visualizzare la capacità degli atomi di combinarsi e permetteva, inoltre, di spiegare, almeno intuitivamente, le proprietà dei composti, soprattutto dei composti organici. La teoria si affermò soprattutto a causa del successo che conseguiva nel razionalizzare i meccanismi delle reazioni organiche a tal punto che viene ancora usata nei testi di chimica. Non riusciva però a spiegare i problemi di stereochimica, l'attività ottica e soprattutto i fenomeni spettrali che stavano assumendo un'importanza fondamentale come metodo per identificare i composti.

La teoria dell'ottetto fu ignorata da molti, soprattutto dai chimici-fisici e dai fisici. In opposizione con il modello di Bohr, modello essenzialmente dinamico, i due chimici americani proponevano un modello statico, interessati com'erano a interpretare i problemi di reattività delle molecole.

Un ragionamento tipico della teoria di Lewis-Langmuir può essere il seguente. Si era osservato che la sostituzione di un atomo di cloro a un atomo di idrogeno in un acido organico aumentava il carattere acido del composto. La spiegazione che veniva data è abbastanza semplice. Il cloro, nel tentativo di completare il suo ottetto, acquista un elettrone e, dato che il suo nucleo contiene piú protoni, quindi è piú positivo, attrae l'ottetto piú vicino, rendendo cosí piú poveri in elettroni gli altri atomi, in modo particolare l'idrogeno, che, avendo un solo elettrone, si dissocierà cosí piú facilmente come ione H^+. Tutti i gruppi della chimica organica potevano essere classificati secondo la capacità di attirare o di cedere gli elettroni ad altri gruppi. Tale effetto poteva essere pensato propagarsi lungo lo scheletro di atomi di carbonio, che forma la struttura base di ogni composto organico.

La teoria di Laurent, cui si è accennato prima, trovava una base di conferma elettronica e non solo sperimentale. Di qui il perdurare nei testi di chimica organica sia della classificazione in base ai radicali, sia della teoria di Lewis.

La nuova teoria del legame riuscí a prevedere anche alcuni fatti sperimentali. Lo ione H^-, previsto dalla teoria in quanto la sua struttura elettronica corrisponde a quella dell'elio, fu messo in evidenza sperimentalmente effettuando l'elettrolisi del composto formato da idrogeno e litio, l'idruro di litio.

Nell'ambito dei chimici organici dell'epoca, la nuova teoria fu accettata subito in Inghilterra, respinta in Germania, quasi ignorata negli Stati Uniti. I chimici americani passavano da un atteggiamento di puro disprezzo – giudicando, come fece il chimico Frey, «metafisiche» tutte le teorie sulla struttura atomica, con i loro ipotetici gusci di elettroni o orbite elettroniche – ad un atteggiamento di attesa. Lo spirito generale era che il cammino piú sicuro per un chimico organico fosse di pensare la valenza elettronica «semplicemente come il numero di cariche negative che un atomo può acquistare o perdere nel raggiungere il mistico numero di otto, lasciando i problemi di struttura atomica ai fisici» [Conant 1928, p. 25].

Nell'opposizione tedesca bisogna vedere due fattori. Il primo, di ordine socio-politico, è l'isolamento scientifico della Germania negli anni dopo la prima guerra mondiale. La crisi economica rendeva praticamente impossibili non solo i viaggi all'estero, ma anche l'abbonarsi alle riviste straniere. Haber, uno dei piú noti chimici dell'epoca, scrive a Langmuir il 23 dicembre 1921 pregandolo di inviargli gli estratti dei suoi ultimi lavori: «Le nostre comunicazioni con gli Stati Uniti sono in una terribile situazione; non mi è stato possibile trovare tale giornale in tutta Berlino» [citato in Kohler 1975]. Nello stesso tempo gli scienziati tedeschi si erano imposti un autoisolamento nei confronti degli anglo-americani. La risposta nazionalista alla sconfitta militare faceva parte dello spirito del dopoguerra e aveva conquistato anche gli ambienti scientifici.

Il secondo motivo è che le due figure principali della teoria della valenza in Germania furono due fisici: Stark e Kossel. La teoria dell'ottetto di Kossel era apparsa quasi contemporaneamente a quella di Lewis. I chimici tedeschi preferirono adottare le idee dei due fisici. Bisogna aggiungere, però, che la teoria elettronica del legame rimase marginale fra i chimici organici tedeschi. Quelli che la adottarono, in genere, venivano da esperienze di ricerca nei campi della chimica-fisica; la maggior parte continuò ad usare le idee tradizionali, assumendo un atteggiamento di chiusura, sicuri della superiorità conquistata dalla chimica tedesca.

In Inghilterra la diffusione della teoria è legata essenzialmente al tentativo di alcuni chimico-fisici di risolvere la contraddizione tra modello dinamico e modello statico dell'atomo. La conseguenza fu che si riuscí a dimostrare l'utilità della teoria nell'interpretare alcuni problemi chimici. In parte contribuí al successo l'abitudine di alcune scuole della chimica organica a cercare una teoria che potesse rendere conto di una reazione organica, spiegandone le singole tappe, i meccanismi di reazione.

In questo senso tutte le teorie, dai legami di Kekule ai tetraedri di van't Hoff, ai legami elettrostatici, erano state utilizzate. La teoria di Lewis-Langmuir aveva il vantaggio di essere semplice e di inglobare molti fatti sperimentali. Questa ricerca affannosa di una teoria che permettesse di fare notevoli progressi nel campo della chimica organica era anche la conseguenza del brusco risveglio che la prima guerra mondiale aveva provocato nella chimica inglese. La ricerca svolta in alcuni centri universitari era competitiva a livello internazionale, ma l'industria chimica era praticamente priva di laboratori di ricerca e

di personale qualificato. La presenza dell'industria chimica inglese sui mercati internazionali era marginale. La maggior parte dei prodotti chimici erano prodotti dalla Germania e dalla Svizzera. Cosí l'Inghilterra, che alla metà del XIX secolo esportava la metà della sua produzione chimica, allo scoppio della prima guerra mondiale importava i nove decimi del fabbisogno nazionale. Anche i coloranti utilizzati per le uniformi militari erano importati dalla Germania.

Per rimediare a questa situazione furono nominate numerose commissioni governative. Tutte furono d'accordo nel riconoscere che i principali rimedi consistevano in un rinnovamento degli studi e in un intervento diretto dello Stato per creare alcuni grossi complessi di produzione, raggruppando le piccole industrie che, bene o male, sopravvivevano. Fu recensito il personale chimico qualificato che lavorava nell'industria. Volendo essere ottimisti, tale personale rappresentava circa un terzo di quello che lavorava nelle industrie tedesche. Inoltre, mentre l'84 per cento dei chimici tedeschi impegnati nell'industria aveva un titolo universitario, seguito da un'esperienza di ricerca nei laboratori universitari, solo il 34 per cento dei chimici inglesi si trovava nella stessa situazione. Di questo 34 per cento, la metà aveva studiato in Germania.

Sull'esempio tedesco furono rinnovati completamente gli studi e lo stato finanziò in modo massiccio la formazione di nuovi laboratori universitari. La Germania, allo scoppio della guerra, spendeva per la sovvenzione dei laboratori chimici una cifra dieci volte superiore a quella inglese.

Nell'opera di rinnovamento, Thomson e Rutherford ebbero una grande influenza, spingendo i chimici ad interessarsi alle nuove teorie del legame.

8. *Sviluppo dell'atomo di Bohr e comparsa di una nuova teoria: la meccanica quantistica.*

Il modello atomico proposto da Bohr si rivelò molto utile nell'interpretare i risultati della spettroscopia atomica, essenzialmente i fenomeni derivanti dall'interazione di una radiazione con gli elettroni. Il modello atomico era un modello planetario. Le equazioni di Hamilton, usate dagli astronomi per descrivere i sistemi celesti con orbite periodiche, vennero applicate all'atomo. Il confronto continuo con i risultati della spettroscopia serviva da guida all'analisi teorica e forniva la terminologia necessaria a classificare le grandezze fondamentali.

Accanto al primo numero quantico introdotto da Bohr per indicare il livello di energia dell'orbita elettronica, Sommerfeld introdusse un secondo e un terzo numero quantico, entrambi funzione del primo. L'interpretazione geometrica data dallo stesso Sommerfeld è che l'orbita è essenzialmente ellittica, simile a quella dei pianeti. Un'ellissi è caratterizzata da due raggi, cioè da due numeri quantici.

Se si osserva lo spettro atomico in presenza di un campo magnetico, esso risulta notevolmente modificato (effetto Zeeman). Dalla fisica classica si conosceva che una carica elettrica in movimento genera un campo magnetico. L'elettrone è una carica elettrica che ruota, dunque è dotato di un momento magnetico, che interagisce con un campo magnetico esterno assumendo valori discreti. Il terzo numero quantico può essere messo in relazione con la posizione che l'ellissi occupa rispetto ad un campo magnetico esterno.

Il quarto numero quantico fu introdotto da Pauli. Il numero quantico di spin rappresenta essenzialmente la rotazione dell'elettrone intorno al suo asse. Goudsmith e Uhlenbeck mostrarono che con queste quattro grandezze si potevano interpretare le caratteristiche degli spettri atomici.

Bohr non aveva mai abbandonato l'idea di poter descrivere, in base alla nuova meccanica, le proprietà degli atomi, cercando d'interpretare la regolarità della tavola periodica degli elementi. Nel 1918 intraprese una esposizione sistematica dei principî e risultati della nuova meccanica atomica, pensando di poter estendere la sua analisi alle proprietà delle molecole. Questa esposizione non fu portata a termine. Nello stesso tempo, guidato solo da considerazioni teoriche, spinse alcuni dei suoi collaboratori a cercare sperimentalmente un elemento che prevedeva dovesse occupare il settantaduesimo posto nella tavola periodica. Bohr stesso annunziò il risultato positivo di questa ricerca nel 1922, durante il discorso di ricezione del premio Nobel. Il nuovo elemento fu chiamato afnio, dal nome latino di Copenhagen.

Incoraggiato da questi successi, Bohr affidò a Pauli, allievo di Sommerfeld, l'incarico d'interpretare i fenomeni spettroscopici lungo la tavola periodica degli elementi. Nel corso della sua analisi, Pauli si accorse che tutto diventava semplice se ad ogni elettrone venivano assegnati in modo unico i quattro numeri quantici. La formulazione di questo principio è conosciuta con il nome di principio di esclusione di Pauli.

Intorno al 1925, mentre i chimici si sentono incoraggiati a proseguire ad usare il loro modello di valenza dai risultati positivi dei fisici, questi ultimi provano un senso di fastidio di fronte al pesante fardello di regole empiriche necessarie per descrivere i fenomeni atomici. In un altro modo si potrebbe dire che mentre i chimici sono soddisfatti di poter descrivere qualitativamente le proprietà delle molecole con un modello atomico, i fisici cercano un modello matematico generalizzabile.

La svolta centrale si ebbe nel 1926.

Il punto di partenza di Heisenberg-Born-Jordan è una critica dei concetti classici di velocità, posizione e traiettoria. L'applicazione di queste grandezze ai sistemi atomici deve basarsi su un criterio di osservazione e di misura. Le grandezze fisiche non devono piú essere rappresentate da funzioni ma da operatori. I valori associati a tali operatori sono dei numeri reali e corrispondono ai risultati numerici della misura di tali grandezze. Il nucleo principale dell'algebra degli operatori è costituito dal principio d'indeterminazione, conseguenza diretta della regola di moltiplicazione fra operatori, e dove la costante di Planck diventa l'unità fondamentale di misura. In base a questo principio, un'esperienza concepita per conoscere esattamente la posizione di un elettrone non permette di conoscere la quantità di moto e viceversa.

Crollava cosí la concezione del determinismo valido nella fisica classica.

Classicamente, infatti, è possibile determinare sperimentalmente ed esattamente sia la posizione sia la velocità di un oggetto. In questo modo la traiettoria è conosciuta. Nella formulazione di Heisenberg-Jordan-Born tale determinazione non è piú possibile a causa della perturbazione del sistema causata dall'osservatore. Per le grandezze macroscopiche tale perturbazione è trascurabile dato il piccolissimo valore della costante di Planck.

Il punto di partenza di De Broglie si basa sulla equivalenza onda-corpuscolo stabilita da Einstein per il fotone, cercando di trasferirla all'elettrone. Schroedinger sviluppò l'ipotesi di partenza di De Broglie, stabilendo che il movimento di un elettrone non deve essere fondato sulle traiettorie classiche, soluzioni delle equazioni di Newton: il moto dell'elettrone è descritto da un'onda associata, la funzione d'onda, soluzione di un'equazione differenziale alle derivate parziali, l'equazione di Schroedinger. Lo stesso anno, Schroedinger mostrò l'equivalenza del formalismo ondulatorio con quello operatoriale. La funzione d'onda, secondo l'interpretazione data da Born, acquista un significato piú astratto. Essa rappresenta lo stato fisico di un sistema e permette il calcolo dei valori medi delle grandezze fisiche che descrivono il sistema in quello stato dato.

Il formalismo di Schroedinger aveva il vantaggio, trattandosi essenzialmente di equazioni differenziali, di essere familiare ai fisici; e le funzioni d'onda potevano essere caratterizzate da numeri interi, reintroducendo cosí i numeri quantici.

I risultati positivi ottenuti nella trattazione dell'atomo d'idrogeno incoraggiarono all'applicazione ad altri sistemi.

9. *Il legame chimico alla luce della meccanica quantistica.*

Bohr aveva considerato la molecola d'idrogeno come un sistema a due nuclei con i due elettroni ruotanti in un'orbita perpendicolare ai due nuclei. Trovare l'equazione di Schroedinger per un tale sistema non è complicato, ma è impossibile risolvere l'equazione stessa. Si doveva ricorrere quindi ad alcune approssimazioni.

Nel formare la molecola, i due atomi d'idrogeno si influenzano a causa delle interazioni reciproche, cioè gli elettroni saranno descritti da una funzione leggermente diversa da quella negli atomi isolati. L'energia della molecola sarà rappresentata da due termini: uno rappresenta l'energia corrispondente all'energia dei due atomi, l'altro deriva essenzialmente dall'interazione dei due atomi fra di loro. Quest'ultimo termine sarà funzione della distanza fra i due nuclei.

Nel descrivere tale sistema si deve tener conto di alcuni fattori: la funzione che descrive la molecola deve avere le stesse proprietà di simmetria della configurazione dei nuclei; si deve tener conto della indistinguibilità degli elettroni, cioè ogni elettrone può appartenere indifferentemente ad ognuno dei due nuclei; il principio di esclusione di Pauli deve essere rispettato.

Basandosi su queste ipotesi, Heitler e London eseguirono un calcolo usando come funzione di partenza per descrivere la molecola il prodotto degli orbitali

dell'atomo d'idrogeno. Facevano variare leggermente le funzioni atomiche fino ad ottenere un minimo per il valore dell'energia del sistema. La distanza tra i due nuclei, distanza di legame, e l'energia erano in buon accordo con i dati sperimentali. Si poteva pensare che un miglioramento nelle funzioni di partenza avrebbe permesso di ottenere dei risultati migliori. La stabilità e l'esistenza della molecola erano legate alla coppia di elettroni che si trovava localizzata ad uguale distanza dai due nuclei, in pratica condivisa dai due nuclei. Si ritrovava cosí l'idea alla base della teoria del legame chimico di Lewis-Langmuir. I chimici organici potevano tranquillamente continuare ad usarla in modo del tutto empirico.

L'ipotesi di partenza di Heitler-London era che il legame nella molecola d'idrogeno non fosse dovuto al trasferimento di un elettrone da un atomo ad un altro. Ma a priori nessuno poteva negare questa possibilità. Slater, che aveva partecipato direttamente alla nascita della meccanica quantistica nel laboratorio di Bohr, e Pauling proposero che la funzione d'onda di una molecola dovesse essere rappresentata dalla combinazione lineare di tutte le funzioni d'onda che potevano essere associate per una molecola a tutte le strutture teoricamente possibili. Cosí per la molecola d'idrogeno si dovevano considerare anche le funzioni associate a strutture come $H^{+-}H^{-}$ e $H^{--}H^{+}$, simmetriche ma diverse dal punto di vista elettronico. La funzione finale è la sola funzione vera, che descrive il sistema e le sue proprietà. Le strutture singole o forme canoniche che si usano per determinare le funzioni di partenza sono semplicemente delle finzioni.

Il problema principale è che la determinazione di tutte le forme canoniche possibili diventa un rompicapo con l'aumentare del numero di atomi presenti. Per lo ione SO_4^{--}, per esempio, le strutture possibili sono 4 600 000, è cioè praticamente impossibile prenderle in considerazione tutte.

Se si considera valida l'ipotesi che il legame sia dovuto essenzialmente agli elettroni piú esterni, gli elettroni interni possono essere considerati praticamente imperturbati, cioè descrivibili con le stesse funzioni usate per l'atomo isolato. Nei composti organici, inoltre, è indifferente vicino a quale atomo di carbonio si trovi l'elettrone. Le funzioni atomiche di partenza sono cosí una buona approssimazione per calcolare gli orbitali molecolari, quelli cioè che appartengono a tutta la molecola, come combinazione lineare degli orbitali atomici. Cosí, a partire da dieci orbitali atomici si otterranno dieci orbitali molecolari. In ogni orbitale sarà possibile mettere due elettroni, rispettando il principio di esclusione di Pauli. Ad ogni orbitale corrisponde una coppia di elettroni, come veniva supposto nella teoria del legame di Lewis-Langmuir.

Con la teoria degli orbitali molecolari si è fatto un passo avanti nella descrizione dei sistemi chimici. Non tutti gli orbitali molecolari saranno occupati dagli elettroni, alcuni restano vuoti. In questo modo è possibile interpretare gli spettri di assorbimento o di emissione. Una transizione sarà dovuta al passaggio di un elettrone da un orbitale occupato ad uno vuoto, esattamente come per gli atomi gli spettri erano dovuti al passaggio di un elettrone da un livello atomico ad un altro. L'energia a cui la transizione avviene sarà uguale alla differenza di energia tra i due orbitali. Il modello serve ad interpretare

i risultati sperimentali, come l'accordo con i dati sperimentali indica la bontà del modello, permettendo cosí di utilizzare la funzione d'onda molecolare ottenuta per prevedere i risultati di altre misure.

La ricerca in fotochimica ha ricevuto uno stimolo considerevole dalla possibilità d'interpretare i risultati sperimentali con il metodo degli orbitali molecolari. Cosí la definizione di molecola è quella di un sistema a piú nuclei con una nube di carica negativa, piú o meno densa nei vari punti dello spazio, ma che si estende su tutta la molecola. In realtà questa descrizione non piace al chimico, abituato dalla strutturistica a vedere i legami come bastoncini che puntano verso determinate direzioni dello spazio. D'altronde le formule di struttura sono utili nella pratica di laboratorio.

L'arbitrarietà con cui è possibile scegliere le funzioni di base per costruire gli orbitali molecolari, permette di scegliere delle funzioni che si sviluppano secondo una direzione privilegiata dello spazio intorno ad un nucleo. Si possono utilizzare sempre gli orbitali atomici come punto di partenza, ma trasformandoli prima, imponendo come condizione che essi siano indistinguibili tra di loro e che rappresentino il massimo di distribuzione della densità elettronica secondo direzioni ben precise dello spazio. Le funzioni molecolari risultanti sono costituite cosí da coppie dei nuovi orbitali, aventi un campo di densità in comune, e la densità elettronica è localizzata tra due nuclei.

In pratica si tratta della trasposizione su basi matematiche della teoria del legame di Lewis-Langmuir, cui è stata associata l'idea del tetraedro di van't Hoff. Gli orbitali atomici trasformati sono chiamati orbitali ibridi e possono essere rappresentati geometricamente come un insieme di orbitali che punta verso i vertici di un tetraedro, un triangolo equilatero, una bipiramide a base quadrata, una bipiramide a base triangolare.

I legami chimici vengono cosí classificati in due tipi: l'uno è localizzato fra due atomi; l'altro, in nessun modo localizzabile fra due atomi, può venire rappresentato solo come esteso su tutta la molecola. I primi vengono chiamati legami semplici, i secondi doppi o tripli, secondo la densità elettronica associata. Nel caso di questi ultimi, il legame non deve essere considerato dotato di individualità propria, ma può essere pensato soltanto nel suo insieme.

Spesso sono classificati sotto la voce di legami alcuni fenomeni dovuti all'interazione degli atomi o delle molecole fra di loro. Si parla cosí di legami di Van der Waals o di legami idrogeno. I primi sono assimilabili piuttosto alle forze d'interazione fra molecole o fra atomi, i secondi sono piú vicini ad un legame vero e proprio.

Analizzando la densità elettronica che risulta da un calcolo di orbitali molecolari, si osserva che alcuni atomi sono circondati da una nube piú densa. Questi atomi sono classificati come elettronegativi e si può stabilire una scala di elettronegatività confrontando gli elementi con l'idrogeno. Tale dissimmetria nella distribuzione della carica elettronica corrisponde ad una grandezza macroscopica, il momento dipolare.

Quando in una molecola è presente un atomo d'idrogeno vicino ad un altro atomo elettronegativo, la densità di carica intorno all'idrogeno è bassa a tal punto

da farlo rassomigliare piú ad un protone che ad un atomo d'idrogeno. Se sono presenti altre molecole simili, esse cercheranno di orientarsi in modo che l'elemento elettronegativo si metterà di fronte all'idrogeno per annullare l'eccesso di carica positiva. Si formano cosí dei legami parziali, molto piú deboli di un legame ordinario, ma responsabili della formazione di una struttura ordinata all'interno del sistema. La presenza di questo legame rende conto delle proprietà delle macromolecole in generale e in particolare di alcune molecole importanti nella biologia molecolare.

10. *Il protone risolve le ultime anomalie della struttura atomica.*

Restava ancora un punto non ben chiaro sulla costituzione degli atomi: il numero atomico, cioè il numero d'ordine occupato dagli elementi nella tavola periodica, non aveva lo stesso valore del peso atomico come ci si sarebbe dovuto attendere dall'uguaglianza tra protoni ed elettroni, il cui numero era rappresentato dal numero atomico. La massa di un atomo si può infatti pensare uguale essenzialmente a quella del nucleo, dato che un elettrone ha una massa uguale a circa 1/2000 della massa di un protone. Il nucleo era considerato costituito da protoni ed elettroni. L'eccesso di protoni sugli elettroni, concentrati nel nucleo, costituiva la carica positiva. Nel 1932 Heisenberg e Majorana mostrarono teoricamente l'inconsistenza di tale modello.

In alcuni esperimenti sia Joliot e Curie sia Boethe e Becker avevano osservato delle radiazioni molto penetranti, che non venivano deviate né da un campo magnetico né da un campo elettrico. Cadwick analizzò in modo sistematico questo tipo di radiazione, mettendone in evidenza la natura corpuscolare e determinando il valore della massa, risultata uguale a quella del protone.

Cosí l'atomo fu pensato costituito essenzialmente da un nucleo, formato da protoni e neutroni, e dagli elettroni. Se le forze con cui gli elettroni interagiscono con il nucleo trovavano la spiegazione nella meccanica quantistica, restavano da spiegare le forze in gioco tra le particelle costituenti il nucleo. Nasceva cosí un nuovo campo d'indagine nella fisica: quello delle particelle elementari.

11. *Conclusioni.*

La storia dello sviluppo del concetto di atomo e di molecola può essere schematizzata in tre stadi, corrispondenti all'affinamento di alcune tecniche sperimentali. Cosí dal visibile si è passati al microscopico, dal microscopico al molecolare, dal molecolare alla dimensione elettronica. A ciascuno di questi tre stadi corrisponde una concezione diversa della materia.

Fino a quando la concezione atomista è rimasta un fatto puramente filosofico, come spiegazione ultima delle diverse forme assunte dalla materia, la scienza chimica è rimasta allo stadio dell'osservazione delle qualità. Le forme, il sapore,

l'odore hanno permesso di classificare e di riconoscere una regolarità nelle proprietà della materia.

L'introduzione della teoria atomica in fisica e in chimica ha radicalmente cambiato la struttura di quest'ultima. La chimica cessa di essere la scienza del qualitativo e diventa la scienza del quantitativo. Prima di Dalton non esisteva un sistema per indicare la costituzione delle sostanze chimiche, se non attraverso le proprietà osservabili. Facendo attenzione ai pesi molecolari e alle reazioni i chimici diventano capaci di stabilire la struttura geometrica delle molecole. Le misure volumetriche e gravimetriche erano state usate fino a quel momento semplicemente per dimostrare che qualcosa era avvenuto. Ora diventano lo strumento per misurare, cioè determinare quantitativamente, quello che succede. Si possono ricavare cosí delle leggi generali basate sulla semplice osservazione dei rapporti ponderali.

La concezione atomistica, pur derivando direttamente dall'atomismo classico, si stacca completamente da esso. Gli atomi sono pensati non solo come delle specie interagenti fra di loro per unione, separazione ed impatto, ma possono interagire anche a distanza esercitando delle forze reciproche. È questo il significato profondo della teoria elettromagnetica di Maxwell.

La forza insita nelle teorie atomiste si esprime anche nel provocare lo sviluppo enorme dei processi tecnologici legati alla lavorazione e trasformazione della materia. L'industria dei coloranti diventa il banco di prova dell'utilizzazione della comprensione teorica dei problemi di sintesi per applicarli ai problemi puramente tecnologici. Lo straordinario sviluppo dell'industria chimica, a partire dal 1860, potrebbe servire come guida per dimostrare la vitalità rappresentata dall'atomismo classico.

Ma la concezione atomista ebbe implicazioni piú vaste. Le leggi chimico-fisiche permettevano di comprendere i fenomeni geologici, la storia geologica. Cosí la storia della Terra diventa il risultato di quelle reazioni chimiche che nella pratica di laboratorio si è appreso ad interpretare.

L'estensione del concetto di molecola come massa a sistemi complessi come le macromolecole segna anche il momento della crisi di tale concetto. Anche delle misure sofisticate, come la determinazione con una ultracentrifuga della distribuzione delle varie masse e la loro determinazione con metodi ottici molto raffinati, diventano impossibili. Resta soltanto il concetto di molecola come struttura complessa di piú atomi. Si cerca cosí, disponendo atomi o gruppi di atomi legati rigidamente tra di loro, di determinare la struttura piú probabile che corrisponda alle proprietà osservate, per analogia con altre strutture ben note. Il qualitativo, che ha sempre accompagnato la chimica nel suo sviluppo, riprende il suo posto accanto al quantitativo.

Alla fine del XIX secolo i problemi che interessano i fisici sono essenzialmente due: la radiazione in tutte le sue forme e la costituzione della materia.

L'unificazione di queste due problematiche, che la teoria dei quanti e la relatività fanno intravedere, cambia il concetto che si ha della materia. L'unificazione dei costituenti ultimi in particelle subatomiche ha introdotto nell'osservazione della materia un ordine di analisi in base ad una struttura organizzativa. La materia viene cosí analizzata a livello subatomico, atomico, molecolare, dove i termini perdono il significato di limite preciso ma servono essenzialmente ad indicare il grado di complessità della struttura esaminata. Con questo non si vuol dire che i concetti classici di atomo e di molecola abbiano perso il loro valore nella pratica.

La teoria della costituzione della materia si pone oggi un problema piú ambizioso: unificare i concetti di materia e di vita, come la teoria atomica ha abolito la distinzione tra organico e inorganico. Infatti un atomo contiene potenzialmente la possibilità di essere un gas, un cristallo, un virus. Cambia essenzialmente il grado di complessità.

Nel tentativo di unificare i concetti di materia e di vita il grado di complessità di organizzazione diventa fondamentale. Un virus, un gene isolati non si comportano in modo diverso dalle altre molecole. Isolabili, cristallizzabili, essi sono analizzabili in termine di legame chimico, di distribuzione nello spazio di nuclei e di densità elettronica. Ma a differenza della maggior parte delle molecole conosciute, inseriti in un organismo piú complesso diventano capaci di assumere delle caratteristiche uniche, formando nell'insieme le proprietà che vengono attribuite ad una cellula.

Il problema che si pone è la ricerca delle relazioni esistenti fra i diversi gradi d'ordine delle strutture della materia.

A Kekule fu domandato un giorno cosa pensasse della esistenza degli atomi. «La domanda se gli atomi esistano o no ha scarso significato dal punto di vista chimico: la sua discussione appartiene alla metafisica. Nella chimica noi dobbiamo decidere se l'assunzione degli atomi sia un'ipotesi adatta alla spiegazione dei fenomeni chimici... Da un punto di vista filosofico, io non credo nell'esistenza degli atomi, prendendo il termine nel suo significato letterale di particelle indivisibili della materia. Mi aspetto piuttosto che si troverà un giorno per quello che noi ora chiamiamo atomi una spiegazione matematico-meccanica, che ci renderà i concetti come peso atomico, atomicità e numerose altre proprietà dei cosiddetti atomi. Come chimico, tuttavia, io guardo l'ipotesi degli atomi non solo come utile, ma come assolutamente necessaria in chimica. Voglio spingermi ancora piú avanti e dichiarare che l'atomo chimico esiste, facendo attenzione a che il termine indichi quelle particelle di materia che non subiscono alcuna divisione nelle trasformazioni chimiche. Se un progresso della scienza conducesse a una teoria sulla costituzione dell'atomo chimico, ciò provocherebbe pochi cambiamenti nella chimica stessa. Gli atomi chimici resteranno sempre l'unità chimica, e, per le considerazioni riguardanti in modo particolare la chimica, noi potremo sempre cominciare dalla costituzione degli atomi e accontentarci delle espressioni semplici che ne risultano, in altre parole dell'ipotesi atomica. Noi possiamo infatti adottare il punto di vista di Dumas e Faraday, che sia che la materia sia atomica o no, questo è certo, che benché essa sia atomica, ci apparirebbe cosí come essa ci appare».

A parte le previsioni sul futuro della teoria atomica, che ci si potrebbe divertire a considerare profetiche o no; a parte la difesa un po' fredda dell'atomismo, comprensibile nel clima acceso di antiatomismo, quello che è interes-

sante, perché valido ancor oggi, è il giudizio della posizione dei chimici nei confronti dello sviluppo della teoria atomica.

I progressi della conoscenza dell'atomo hanno toccato marginalmente i chimici nella pratica di laboratorio, anche se oggi nessuno piú dubita della realtà dell'atomo. Cosí la pratica corrente del chimico si può analizzare in base allo sviluppo del concetto di molecola.

Nel lavoro di sintesi, anche se altamente sofisticato, i concetti di massa e di struttura sono i soli in gioco. Che essi bastino a guidare il chimico è dimostrato anche dal fatto che sulle leggi derivate da queste due caratteristiche è possibile eseguire la previsione del decorso di una reazione di sintesi. Anzi, tali leggi, basate per lo piú su osservazioni empiriche, sono talmente ben sperimentate che è possibile programmarle su calcolatore, ottenendo cosí automaticamente le previsioni sulla possibilità di una sintesi.

Bisogna notare che il concetto di struttura, come distribuzione in sequenza ordinata di atomi, è difficile da abbandonare, in quanto il nome che il chimico dà ai suoi composti si basa sulla sequenza degli atomi nella molecola. Il problema non è del tutto semplice. Le molecole conosciute sono circa quattro milioni, cioè piú delle parole contenute in qualsiasi dizionario. Classificare e ordinare tale massa di composti e dar loro un nome razionale, cioè che nello stesso tempo contenga alcune delle informazioni indispensabili a caratterizzare il composto, non è impresa da poco. D'altronde la ricerca chimica non è terminata. Qualsiasi sistema concepito per chiamare i composti deve essere talmente flessibile da poter inserire qualsiasi nuovo composto che potrà essere sintetizzato. La ricerca di un modo unico sufficientemente generalizzato e malleabile viene condotta con l'aiuto dei calcolatori che possono memorizzare facilmente tale massa di risultati. Non si è ancora riusciti a trovare un metodo soddisfacente, anche se ne esistono alcuni nella pratica corrente.

La molecola contiene anche delle informazioni sulle qualità che caratterizzano il suo comportamento. Il chimico-fisico si trova spesso di fronte al compito di determinare attraverso le leggi della fisica le proprietà che si possono attribuire ad una classe di composti. Generalmente ciò accade in vista di poter stabilire delle regole a priori della reattività di un composto in base, per esempio, all'energia associata ad un legame, alla struttura di una molecola, al suo comportamento in soluzione. In questo senso egli si accosta al lavoro del chimico che s'interessa dei meccanismi di reazione, e utilizza gli stessi metodi del chimico-fisico per cercare di determinare le singole tappe di una reazione. Entrambi si occupano dell'aspetto qualitativo dell'unica grandezza quantitativa rappresentata dalla molecola. Nel loro lavoro sono spesso aiutati dai chimici teorici che cercano, con l'aiuto delle teorie del legame chimico, di districarli nell'analisi dei risultati sperimentali in rapporto alla complessità della struttura molecolare.

Tale distinzione si riflette anche a livello degli studi. Il chimico viene addestrato prima a classificare in base alle proprietà (analisi qualitativa), e poi a determinare i rapporti di combinazione (analisi quantitativa). Passa cosí allo studio dei meccanismi di reazione e della chimica fisica, cioè allo studio del qualitativo

sofisticato e dell'organizzazione di queste informazioni per analizzare la reattività. Solo alla fine, seguendo qualche corso facoltativo, potrà apprendere i metodi sofisticati del legame chimico e come la problematica chimica si colleghi alla problematica piú generale dell'unità della struttura della materia.

In questo senso gli studi e la pratica sono organizzati secondo il punto di vista che, forse a maggior ragione, difendeva Kekule. Il problema che si pone è quello dell'unificazione epistemologica in funzione delle ultime conoscenze: l'unità della costituzione della materia secondo livelli di organizzazione a complessità crescente. [G. R.].

Avogadro, A.
1811 *Essai d'une manière de déterminer les masses relatives des molécules élémentaires des corps, et les proportions selon lesquelles elles entrent dans ces combinations*, in «Journal de Physique», LXXIII, pp. 58-76.

Bohr, N.
1913 *On the constitution of atoms and molecules*, in «Philosophical Magazine», XXVI, pp. 1-25, 476-502, 857-75; ora in L. Rosenfeld, *On the Constitution of Atoms and Molecules. Papers of 1913 reprinted from the Philosophical Magazine*, Munksgaard, København-Benjamin, New York 1963.

Carnot, S.
1824 *Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance*, Bachelier, Paris.

Conant, J. B.
1928 *Atoms, molecules, ions*, in «Journal of Chemical Education», V.

Dalton, J.
[1802] *A New System of Chemical Philosophy*, Bickerstaff, London 1808-27.

Dulong, P.-L., e Petit, T.-A.
1819 *Recherche sur quelques points importants de la théorie de la chaleur*, in «Annales de Chimie et de Physique», X, pp. 395-413.

Einstein, A.
1905 *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt*, in «Annalen der Physik», XVII, pp. 132-48.
1907 *Die Plancksche Theorie der Strahlung und die Theorie der spezifischen Wärme*, in «Annalen der Physik», XXII, pp. 180-90.

Hermann, A.
1971 *The Genesis of the Quantum Theory (1899-1913)*, Mit Press, Cambridge Mass.

Kekule von Stradonitz, F. A.
1859 *Über die Konstitution und die Metamorphosen der chemischen Verbindung und über die chemische Natur des Kohlenstoff*, in «Annalen», CVI, pp. 129-52.

Kohler, R. E. jr
1975 *The Lewis-Langmuir Theory of Valence and the Chemical Community, 1920-1928*, in R. McCormmach (a cura di), *Historical Studies in the Physical Sciences*, V, Princeton University Press, Princeton N.J.

Koyré, A.
1955 *Tricentenaire de Gassendi*, in *Actes du Congrès de Digne, 1955*, Presses Universitaires de France, Paris 1957.

Partington, J. R.
1964 *A History of Chemistry*, IV, Macmillan, London.

Thomson, J. J.
1899 *On the masses of the ions in gases at low pressures*, in «Philosophical Magazine», XLVIII, pp. 547-67.
1907 *The Corpuscular Theory of Matter*, Constable, London.
Van't Hoff, J. H.
1875 *La chimie dans l'espace*, Bazendijk, Rotterdam (2ª ed. 1887).

L'evoluzione dei concetti di atomo e molecola è legata al progressivo perfezionamento delle tecniche d'indagine (cfr. **esperimento**) e si accompagna al mutare dei criteri d'interpretazione della **materia**. La **fisica** contemporanea, grazie alla teoria della **relatività** e all'applicazione dei **modelli** e delle **strutture matematiche** della meccanica quantistica (cfr. **quanta**), ha fatto definitivamente tramontare la concezione atomistica classica, incentrata sugli aspetti speculativi della **natura** e del **mondo** (cfr. anche **elementi**, **alchimia**), e ha portato al superamento del concetto di molecola intesa come «oggetto» dotato di definita struttura spaziale e proprietà misurabili. Secondo le nuove teorie della costituzione della materia i termini 'subatomico', 'atomico', 'molecolare' perdono infatti la connotazione di limiti precisi. Essi servono a indicare semplicemente un diverso grado di complessità (cfr. **semplice/complesso**) di **organizzazione** della **struttura** esaminata, e si confondono nel concetto di **particelle** di un **sistema** in equilibrio (cfr. **equilibrio/squilibrio**) interagenti (cfr. **interazione**) in base a determinati contenuti di **energia**. L'obiettivo ultimo, dopo l'eliminazione dell'apparente dicotomia **organico/inorganico**, resta l'unificazione dei concetti di materia e di **vita**, ossia la spiegazione delle **cellule** e degli **organismi** viventi in termini di distribuzione di nuclei e di densità elettronica.

Conservazione/invarianza

1. *Introduzione.*

Le leggi di conservazione nelle scienze fisiche sono le leggi o principî che stabiliscono quali proprietà di un sistema isolato rimangono costanti allorché il sistema si sviluppa nel tempo. In altre parole, esse affermano che, mentre il sistema subisce un cambiamento, alcune delle sue proprietà non variano. Queste proprietà sono talvolta definite *invarianti* del sistema o costanti di moto.

È da osservare immediatamente che i problemi della conservazione sono legati a uno dei problemi basilari della scienza e della filosofia, forse a quello fondamentale: la comprensione del mutamento. I principî che enucleano gli elementi immutevoli nel mutevole universo sono oggi considerati i piú generali e importanti nella gerarchia delle leggi scientifiche.

Analizzando le varie trasformazioni che subisce un sistema, può risultare che alcune di esse hanno una natura speciale in quanto lasciano inalterati tutti i fattori «interessanti». Ad esempio si potrebbe essere tentati di affermare che la realizzazione dello stesso esperimento in un punto diverso nello spazio non muterà alcun fattore significativo e porterà quindi alle medesime conclusioni: o, meglio ancora, che le stesse leggi regolano entrambi gli esperimenti nonostante la variazione di posizione nello spazio. Si potrebbe quindi dire che le leggi della natura sono invarianti per una trasformazione di questo tipo.

Il problema potrebbe anche essere formulato in termini di simmetrie. La simmetria di un oggetto, secondo uno dei significati del termine, è un'operazione che porta l'oggetto in uno stato indistinguibile da quello originario. Ad esempio la rotazione di un cerchio è un'operazione di questo genere, poiché il cerchio ruotato non è distinguibile da quello primitivo. Per analogia si può parlare di operazioni di simmetria in relazione a leggi scientifiche: sono operazioni, su sistemi fisici o su sistemi di coordinate, che non modificano le leggi, ossia operazioni che lasciano le leggi invarianti.

Le leggi di conservazione individuano proprietà «privilegiate», mentre le considerazioni sulla simmetria discriminano operazioni «privilegiate». I due concetti sono strettamente vincolati nelle scienze fisiche. Le considerazioni sulla simmetria portano alla scoperta delle leggi di conservazione e la scoperta delle proprietà conservate indica l'esistenza di simmetrie incorporate.

Il termine 'invarianza' è talvolta usato in altri contesti ove ha significati diversi da quello che gli si è qui attribuito. Per citare soltanto il piú importante, nel contesto biologico la parola 'invarianza' di solito si riferisce all'invarianza riproduttiva, ossia alla capacità dell'organismo di replicare se stesso in modo da preservare la struttura specifica della specie. Una descrizione del meccanismo di duplicazione si può trovare altrove, e quindi non vi è alcuna necessità di discutere l'invarianza riproduttiva in questa sede. Anche le simmetrie sono qui trattate soltanto in relazione alla conservazione e all'invarianza.

Questo articolo si occupa di tre argomenti: 1) leggi di conservazione classiche; 2) leggi di conservazione e invarianza nella fisica contemporanea; 3) aspetti filosofici attinenti alla conservazione e all'invarianza.

Per leggi di conservazione classiche s'intendono le leggi di conservazione della materia, del momento lineare e angolare, della carica elettrica e dell'energia. Nel processo della scoperta e formulazione di queste leggi, l'intuizione, la sperimentazione e le considerazioni metafisiche hanno svolto un loro ruolo. Tuttavia l'importanza relativa di queste considerazioni fu diversa nei vari casi. Il processo della scoperta delle leggi suddette coincise inoltre con l'analisi dei concetti di materia, momento, carica ed energia. Il contemporaneo emergere di concetti e leggi è trattato nel § 2.

La fisica del XX secolo mutò radicalmente il panorama classico: *a*) sono state scoperte nuove leggi di conservazione e sono state rifiutate o modificate alcune delle leggi classiche; *b*) sono stati inventati nuovi formalismi matematici che consentono la derivazione delle leggi di conservazione in modo elegante da semplici assunti; *c*) le leggi di conservazione hanno nella fisica moderna un ruolo metodologico diverso da quello che rivestivano di norma nella fisica classica. Tali punti sono elaborati nel § 3.

Parecchi problemi piú generali sono collegati al nostro tema: *a*) qual è la connessione tra leggi di conservazione e principio di causalità? *b*) le leggi di conservazione sono a posteriori come le altre leggi scientifiche, o sono forse affermazioni aprioristiche? *c*) come nascono e qual è il ruolo dei concetti di conservazione nello sviluppo intellettuale infantile? Tali problemi sono discussi nel § 4.

2. *Leggi di conservazione classiche.*

Osservando il mondo che ci circonda ci s'imbatte in vari tipi di mutamento e in vari tipi di persistenza: la roccia che vediamo sulla spiaggia potrebbe essere stata là per migliaia di anni, e difficilmente vi si può rilevare un qualsiasi cambiamento in un'intera vita di osservazione. Anche i tavoli e le sedie, o le case in cui viviamo, hanno una certa persistenza nel tempo, benché in questo caso i mutamenti siano piú evidenti. Un pezzo di legno brucia: scompare o semplicemente varia la sua forma? Il ghiaccio fonde in acqua, ma lo si può ottenere di nuovo congelando quest'ultima. È lo stesso ghiaccio di prima? Un pendolo oscilla: perché continua a muoversi? Perché dopo un certo tempo perviene a un punto d'arresto? Qual è l'origine del suo movimento? Mi sfrego le mani, ed esse diventano piú calde: da dove proviene il calore? Come mai scompare dopo un breve periodo?

Ai Greci, che erano manifestamente perplessi di fronte a simili questioni, si offrivano due soluzioni estreme: *a*) l'opinione di Parmenide, secondo cui nulla in realtà cambia, al livello infimo delle cose sensibili tutto è costante. Il mutamento è illusorio. *b*) Il punto di vista eracliteo-cratileo, secondo il quale tutto cambia continuamente, ogni cosa è in divenire, nulla mai persiste. La persistenza è illusoria. Entrambe le soluzioni potrebbero essere attaccate sullo stesso ter-

reno, in quanto non spiegavano l'origine dell'illusione (del mutamento o della persistenza). Si rendeva quindi necessaria una sintesi che spiegasse i due fenomeni. Nel corso della storia della scienza sono stati compiuti vari tentativi nel senso di una sintesi del genere. L'essenza di tutte queste sintesi è la teoria secondo cui soltanto certi elementi sono costanti, mentre altri non lo sono. In effetti i cambiamenti si verificano, ma soltanto quelli che non interferiscono con la conservazione degli elementi costanti. Il mutamento è controllato. Il problema cruciale è naturalmente quali elementi siano costanti e quali siano le variazioni legittime; ed è appunto a questo riguardo che si diversificano i vari tentativi.

Le leggi di conservazione discusse in precedenza sono parte della risposta a tale problema.

2.1. Conservazione della materia.

Il problema del cambiamento era strettamente legato alla ricerca di una teoria della materia. Una delle primissime teorie che c'interessa è quella di Empedocle. Secondo lui tutta la materia è composta di quattro elementi che non si possono creare né annichilare. I mutamenti nella costituzione degli oggetti materiali sono spiegati in base a variazioni nei rapporti degli elementi costituenti. Pur trovandoci in presenza del vero e proprio inizio della nozione di costanza della materia, associata a una spiegazione della sua diversità, sarebbe avventato attribuire a Empedocle la legge di conservazione della materia. L'asserzione che la materia non può essere distrutta significa che ciascuna delle sue caratteristiche è costante? Ciò sarebbe del tutto contrario all'esperienza. Si sarebbe tentati di affermare che si conserva la quantità di materia, ma tale concetto dev'essere definito e nulla in tal senso si può rintracciare negli scritti di Empedocle.

Un concetto piú chiaro della costanza della materia si può trarre dalle teorie atomistiche di Leucippo, Democrito ed Epicuro, cosí splendidamente rappresentate da Lucrezio nel *De rerum natura*. Benché compaia nuovamente la formula fondamentale «Nil posse creari de nihilo», essa ha un significato piú preciso in considerazione del modello cui si applica. Le unità ultime di cui la materia è composta – gli atomi – sono indistruttibili e indivisibili. Essi hanno forma e peso, ma sono privi di altre proprietà della materia, quali il colore, il gusto, ecc. Tutte le caratteristiche materiali e tutti i mutamenti sono riducibili a differenze nella configurazione degli atomi. Il numero degli atomi è fisso ma, ciò che piú importa, il peso come proprietà fondamentale degli atomi può fornire una misura della quantità di materia. L'idea che sia il peso a caratterizzare la materia, pur forse banale per il lettore moderno, non è affatto tale nel suo contesto storico. Per Aristotele e la lunga tradizione dei suoi seguaci il peso è una qualità accidentale della materia. Inoltre, la materia non è necessariamente pesante; il fuoco e l'aria sono leggeri, ovvero si suppone che si allontanino dal centro della Terra. Anche molto piú tardi, negli scritti degli alchimisti, sino al XVIII secolo, è rintracciabile l'opinione che la trasformazione di un metallo possa modificarne il peso. La vaga convinzione che vi sia qualcosa di costante nella materia non rimase in esclusiva ai sostenitori della dottrina atomistica, ma è il contenuto che essi

dettero alle loro convinzioni a giustificarne la fama quali precursori della legge di conservazione della materia.

Molti secoli separano queste prime anticipazioni della legge dalla sua moderna formulazione, la quale stabilisce che la massa totale in un sistema chiuso rimane costante attraverso qualsiasi modificazione subita dal sistema. Fu Lavoisier a considerare tale affermazione come una legge scientifica e ad assumersi il compito di darle una convalida sperimentale. Una condizione necessaria per quello sviluppo fu l'ulteriore delucidazione del concetto di materia sia dal punto di vista fisico che da quello chimico.

La meccanica del XVII secolo elaborò il concetto di massa come qualità essenziale della materia e misura della sua quantità. In senso stretto si hanno due concetti di massa: *a*) la massa gravitazionale di un corpo che determina la forza di gravità tra esso e gli altri corpi; *b*) la massa inerziale di un corpo che ne determina la resistenza alla variazione di moto, l'inerzia. Piú precisamente, si tratta del rapporto tra la forza agente sul corpo e l'accelerazione che esso acquista in conseguenza di quella forza. Tuttavia le due masse furono ritenute uguali. Il peso di un corpo è il prodotto della massa per l'accelerazione gravitazionale e varierà quindi in funzione dell'accelerazione, ma, una volta conosciute e tenute in debito conto queste variazioni, il peso – essendo proporzionale alla massa – può sempre fornire una misura della quantità di materia.

Contemporaneamente anche la chimica, seppure da un punto di vista moderno molto meno sviluppata della fisica, contribuí ad accrescere le conoscenze sulla materia. I gas furono considerati materia e si iniziò a studiarne il comportamento. Si cominciò ad applicare le teorie atomistiche alla ricerca chimica e presero foma i concetti di elemento, composto e miscela, anche se non era disponibile alcun metodo per distinguerli in pratica.

La formulazione, da parte di Lavoisier, della legge di conservazione della materia è il risultato del lavoro di tutta una vita sull'analisi chimica. In certi casi la legge è un'ipotesi da dimostrare, in altri un presupposto fondamentale che influenza la ricerca. Il metodo da lui usato consisteva in rigorose misure dei pesi delle sostanze reagenti prima e dopo i processi studiati, processi che andavano dalla combustione all'ossidazione, alla variazione di stato aggregazionale, e a molti altri. Dal nostro punto di vista tre aspetti dell'opera di Lavoisier sono importanti: *a*) tutta la materia è pesante. La sua attenta analisi della combustione sconfisse la teoria del flogisto, secondo la quale la combustione è liberazione di flogisto: una sostanza con eventuale peso negativo; *b*) la materia non può andare perduta in un sistema chiuso: il peso totale è identico prima e dopo ogni processo; *c*) il nuovo metodo, l'analisi quantitativa, consente una comprensione della struttura della materia. Utilizzando tale metodo Lavoisier pervenne alla prima lista di elementi in quanto distinti dai composti. L'aspetto *a*) è in realtà una condizione necessaria per la scoperta della legge, mentre l'aspetto *c*) è già una conseguenza della formulazione della legge. È evidente che le misure del peso sarebbero futili se avessero potuto darsi perdite di peso in un processo chimico.

La legge di conservazione della materia fu il risultato di una piú profonda conoscenza della materia, ma al tempo stesso rese possibile un'ulteriore deluci-

dazione del concetto grazie alle sue applicazioni in chimica. Questo duplice disegno nella scoperta della legge è tipico e si ripeterà nello sviluppo di altre leggi di conservazione.

2.2. Conservazione del momento.

Lo studio del moto fornisce un altro esempio di sintesi tra mutamento e persistenza: il moto è cambiamento, ma è soggetto a rigorose leggi di conservazione che limitano la portata del mutamento possibile. Per quanto concerne le teorie del moto esiste una spaccatura tra «antichi» e «moderni». L'idea del moto uniforme continuo nello spazio vuoto, che costituisce l'autentica base della meccanica del XVII secolo, era assolutamente assurda per Aristotele, a tal punto che egli se ne serví nel quadro di una riduzione all'assurdo per dimostrare l'impossibilità del vuoto. Gli aristotelici, sviati dall'esperienza quotidiana, credevano che ogni movimento richiedesse un agente che lo conservasse. Nel caso problematico di una freccia in volo si supponeva che l'aria fornisse l'agente. Soltanto nel XVII secolo il concetto di moto si liberò da questo fraintendimento, e soltanto allora furono scoperte le leggi di conservazione relative al movimento. Naturalmente, anche la scienza del XVII secolo aveva radici piú profonde: per citare soltanto qualche nome, Guglielmo di Ockham e il suo allievo Buridano dell'Università di Parigi ipotizzarono che un oggetto in movimento possieda «impeto» – un'entità immateriale proporzionale al suo peso e alla sua velocità, che ne determina lo spostamento continuo. Nel XVI secolo anche Benedetti sviluppò una teoria dell'impeto che ebbe grande influsso su Galileo. Tuttavia non vi è dubbio che i maggiori contributi alla comprensione del moto furono portati nel secolo XVII da Galileo, Descartes, Huygens, Newton e Leibniz.

L'opera galileiana *Discorsi e dimostrazioni matematiche intorno a due nuove scienze* è un confronto tra la vecchia teoria aristotelica e la sua nuova scienza della meccanica. Qui egli presenta il principio d'inerzia, secondo il quale un corpo può variare la propria velocità soltanto sotto l'influenza di una forza. Un corpo libero rimarrà sempre in quiete, oppure continuerà a muoversi con la velocità iniziale nella stessa direzione. Uno studio dei lavori di Galileo mostra quanto lungo fu, anche per lui, il processo di abbandono della visione aristotelica della necessità di una forza per provocare il moto uniforme. Alla fine egli rimase aristotelico nel senso che il moto circolare aveva per lui uno status speciale e i moti terrestri differivano da quelli celesti. Egli attribuí quindi erroneamente il moto circolare dei pianeti alla loro inerzia.

I residui dell'aristotelismo furono spazzati via da Descartes. Il principio d'inerzia di Descartes si applica imparzialmente a tutta la materia, e per lui il moto uniforme è soltanto un movimento in linea retta; soltanto per inerzia un pianeta non percorre un circolo. Descartes estese inoltre il principio a comprendere le interazioni tra i corpi, quali le collisioni, e postulò la conservazione di una quantità che talvolta definí quantità di moto. Descartes dette la massima importanza a questa legge di conservazione e tentò di farla discendere da considerazioni metafisiche generali, come l'immutabilità di Dio, che gli pareva evidente. Dio

creò la materia con movimento e quiete nelle sue parti, afferma Descartes, e, poiché l'ordine dell'universo dev'essere costante quanto lo stesso creatore, è impossibile che la quantità di moto originariamente immessa nel mondo possa mai cambiare. Ogni aumento in una qualunque parte dev'essere compensato da una perdita in qualche altra, sicché la quantità totale di moto nell'universo rimane costante.

Descartes non riuscí a formulare una definizione chiara di questa «quantità di moto» (in varie occasioni le diede anche il nome *vis*, termine piuttosto disorientante in quanto Descartes insisteva nel dire che le forze non trovavano posto nella sua fisica). È opinione comune che Descartes avesse in mente quello che definiamo *momento* (il prodotto della massa per la velocità), ma attribuirgli una simile definizione è impossibile perché *a*) il momento è un vettore mentre Descartes si riferiva a uno scalare, e *b*) il concetto di massa venne chiarito soltanto da Newton.

Questa situazione è tipica del processo di formazione dei concetti nella scienza. I concetti si sviluppano attraverso ripetuti tentativi di integrarli in una teoria scientifica. Nel corso di questo processo essi si trasformano da entità vaghe, indeterminate, in concetti scientifici ben definiti. In altre parole sono concetti in divenire. La legge di Descartes richiamò l'attenzione sulla «quantità di moto», suscitando tentativi per spiegarla, i quali a loro volta portarono a un miglioramento della legge originaria. Non vi fu mai un chiaro concetto di momento fino alla scoperta di una legge di conservazione: la scoperta della legge e la formazione del concetto furono una cosa sola.

Quella che è oggi nota come legge di conservazione del momento lineare è un semplice corollario della seconda e della terza legge del moto di Newton, ma Newton non enunciò mai tale corollario come legge di conservazione. L'investigazione di entità conservate non faceva parte del suo programma di ricerca. Descartes, d'altra parte, con la sua convinzione che una legge di conservazione sia implicita nella teoria del moto, giunse soltanto a una formulazione abbastanza imprecisa e a una giustificazione alquanto teologica. Abbastanza simile fu il caso di Leibniz, che condivise l'opinione di Descartes ma postulò la conservazione di un'altra entità, che definí *vis viva*, uno scalare piú tardi chiamato energia cinetica (cfr. oltre, § 2.5). Huygens contribuí alla scoperta di entrambe le leggi di conservazione: egli si rese conto che la quantità conservata di Descartes è uno scalare mentre in realtà ciò che si conserva è un vettore, e studiò anche le collisioni di corpi elastici, pervenendo cosí alla conservazione della quantità denominata da Leibniz *vis viva*.

È importante comprendere che ognuno di questi contributi ebbe la sua influenza sullo sviluppo della fisica. La cospicua posizione delle leggi di conservazione nella nostra immagine della scienza è l'eredità della tradizione cartesiano-leibniziana, ma l'incorporazione delle leggi di conservazione scoperte da Descartes, Huygens e Leibniz in un sistema teorico coerente non avrebbe potuto realizzarsi senza la meccanica newtoniana.

Stando cosí le cose, è meglio presentare la legge in termini di meccanica newtoniana. La seconda legge del moto di Newton per una particella esprime la

forza agente sulla particella in funzione della velocità di variazione del suo momento lineare:

$$\mathbf{F}=\frac{d\mathbf{P}}{dt}$$

ove **F** è la forza, **P** il momento e t il tempo.

Il momento lineare **P** è il prodotto della massa m della particella per la sua velocità **V**. Ne segue immediatamente che in assenza di forza non si ha variazione del momento lineare, ovvero – in altre parole – che il momento lineare è conservato. Poiché si tratta di una legge vettoriale, il momento verrà conservato lungo ognuno degli assi: cioè se F_x è zero, P_x sarà conservato anche se sulla particella agiscono forze F_y e F_z.

Si può generalizzare la legge a sistemi di molte particelle. In tal caso le forze interne tra le particelle debbono essere distinte dalle forze esterne che agiscono sul sistema. La legge di Newton si legge allora: la somma totale delle forze esterne addizionata alla somma totale delle forze interne è pari alla somma delle variazioni dei momenti. Tuttavia, secondo la terza legge del moto di Newton, la legge di azione e reazione, le forze che due particelle esercitano l'una sull'altra sono uguali ma di segno opposto e si trovano sulla congiungente le particelle. Ne consegue che la somma totale delle forze interne al sistema si annulla. Allora la variazione del momento lineare è determinata soltanto dalle forze esterne e, quando queste sono uguali a zero, il momento lineare totale del sistema è conservato. Un modo diverso di rappresentare la legge consiste nell'affermare che le forze interne non influenzano il moto del baricentro di un sistema. Qualsiasi cambiamento del momento in un punto del sistema sarà quindi compensato in un altro, sicché il baricentro proseguirà nel suo movimento iniziale. Molti esempi familiari illustrano la legge: il rinculo di un cannone è dovuto al momento da questi acquisito per controbilanciare il momento del proiettile. Analogamente, un razzo avanza per equilibrare il momento dei gas di scarico.

È ovvio che per una particella libera di massa costante la legge di conservazione del momento lineare non è altro che il principio d'inerzia; tuttavia, la differenza è significativa. In primo luogo, la legge di conservazione del momento si applica a una piú vasta gamma di casi, come ad esempio sistemi con forze interne o sistemi che durante il movimento perdono o acquistano massa. In secondo luogo, il momento si dimostra concetto piú fondamentale sia della massa che della velocità, sicché, anche se quelle quantità non caratterizzano piú il sistema, come avviene nella meccanica quantistica o nel campo relativistico, il momento è sempre un'entità fondamentale che obbedisce a una legge di conservazione.

2.3. Conservazione del momento angolare.

Ogni moto si può rappresentare come combinazione di rotazione e traslazione. Sinora ci siamo occupati soltanto della traslazione. Esiste una legge analoga per il moto rotatorio, la legge di conservazione del momento angolare. Anche qui la legge ha le sue fondamenta nella meccanica newtoniana, ma fu in realtà scoperta soltanto verso la metà del XVIII secolo da Eulero e Bernoulli.

L'equazione, analoga alla seconda legge di Newton per una particella in moto rotatorio, uguaglia il momento meccanico **N** alla velocità di variazione del momento angolare **L**:

$$\mathbf{N}=\frac{d\mathbf{L}}{dt}.$$

N e **L** sono vettori definiti come segue:

Il momento angolare di una particella rispetto a un punto O è il prodotto vettoriale del raggio vettore da O alla particella, **r**, per il momento lineare della particella, **P**:

$$\mathbf{L}=\mathbf{r}\wedge\mathbf{P}.$$

Il momento meccanico **N** rispetto a un punto O è il prodotto vettoriale di **r** per la forza **F**:

$$\mathbf{N}=\mathbf{r}\wedge\mathbf{F}.$$

Ne deriva immediatamente la legge di conservazione: in assenza di momento meccanico il momento angolare della particella si conserva. La legge viene ancora generalizzata per sistemi di molte particelle sommando tutte le particelle e servendosi della terza legge di Newton per dimostrare che la somma dei momenti interni si annulla. In tal modo la variazione del momento angolare è influenzata soltanto dal momento meccanico esterno e, se quest'ultimo vale zero, il momento angolare totale del sistema si conserva.

Il danzatore classico fornisce un buon esempio: se tiene le braccia piú vicine al corpo (riducendo cosí l'**r** di una parte del corpo), egli ruoterà piú velocemente in modo da non perdere momento angolare. Un altro esempio è dato da un giroscopio, che conserva la sua direzione nello spazio indipendentemente dal movimento del veicolo che lo trasporta.

Le leggi di conservazione, nei termini con cui sono qui rappresentate, hanno validità ristretta per parecchi motivi: *a*) la legge di azione e reazione è in molti casi violata, ad esempio nei sistemi di cariche in moto. Si può dimostrare che in tali casi non sono conservati i momenti meccanici qui trattati, ma la loro somma con un altro fattore, quale ad esempio il momento lineare elettromagnetico; *b*) nel campo relativistico il momento dev'essere definito in modo un po' diverso; *c*) i concetti debbono essere generalizzati nella meccanica quantistica. Ad esempio, una particella con massa zero avrà momento nullo in base alla definizione espressa, ma un fotone, la cui massa è zero, è dotato di momento secondo la meccanica quantistica.

2.4. Conservazione della carica elettrica.

In opposizione alle teorie della materia e a quelle del moto che, come si è visto, hanno una lunga storia, la teoria dell'elettricità è relativamente nuova. Nonostante che il fenomeno dell'elettrizzazione di un pezzo d'ambra fosse già familiare ai Greci, ben poco si seppe dell'elettricità sino al XVIII secolo. Forse questa mancanza di conoscenza scientifica era in qualche modo determinata dalle

nostre limitazioni come esseri umani: ogni fanciullo ha un certo concetto della materia e del moto e i bambini possono persino arrivare a una vaga idea di una legge di conservazione per la materia (cfr. § 4.3). Poiché noi stessi siamo corpi in movimento, abbiamo una concezione intuitiva del mondo nella quale la materia e il movimento sono in qualche modo concetti primari. Non altrettanto si può dire per il concetto di elettricità.

Uno slogan ben noto afferma che la scienza non si occupa del «che cosa» ma piuttosto del «come». In altre parole, per conseguire una conoscenza scientifica di un concetto, non abbiamo altra via che lo studio del comportamento dei sistemi cui esso si applica. Poiché è soltanto uno slogan, non lo si deve ritenere di validità generale (quando è possibile una riduzione ad altri concetti si acquisisce una certa conoscenza sul «che cosa») ma, per quanto riguarda i concetti fondamentali, lo slogan è una descrizione realistica. Ad esempio, la teoria gravitazionale di Newton non dice cos'è la gravitazione, ma come si manifesta. Analogamente, non si sa nulla dell'elettricità a parte la conoscenza del comportamento dei corpi che la «trasportano». È piú facile rendersi conto che questa è una descrizione adeguata in merito a concetti dei quali non abbiamo una comprensione intuitiva, che crearsi l'illusione di un sapere essenziale piuttosto che fenomenologico. Se l'interrogativo è «che cos'è la carica elettrica?», oggi se ne sa quanto trecento anni or sono, ma per quanto riguarda il «come» certamente se ne sa di piú, come attesta la tecnologia moderna. È perciò evidente che la scoperta della legge di conservazione della carica rappresenta al tempo stesso un chiarimento del concetto di carica elettrica. Si è descritto il «divenire» dei concetti di materia e moto. Il caso della carica elettrica è forse ancora piú convincente.

Il padre della teoria dell'elettricità è Benjamin Franklin. Egli postulò l'esistenza di una sostanza elettrica, costituita da atomi abbastanza piccoli da penetrare ogni altra materia. L'analogia con gli atomi di materia implicava un'analoga legge di conservazione. Come si è osservato, la legge atomistica di conservazione non era ancora rigorosa e la sua essenza era l'indistruttibilità della materia. La stessa cosa valeva nel caso della legge di conservazione di Franklin per la sostanza elettrica. Gli atomi componenti la sostanza elettrica nella teoria di Franklin sono di *un solo tipo* e, quando sono distribuiti in modo uniforme, non è possibile scoprirli. I fenomeni elettrici derivano dall'accumulo di questi atomi – eventualità che Franklin denominò «elettricità positiva» – o da una loro deficienza – «elettricità negativa». Quando un corpo positivo viene a contatto con uno negativo si verifica una ridistribuzione degli atomi elettrici in modo che le cariche positive e negative *si neutralizzano*. L'intera quantità di elettricità perduta dalla massa positiva in questo processo viene acquistata dal corpo negativo senza perdite. Questa è la massima approssimazione di Franklin alla legge di conservazione della carica elettrica, ma egli non si serví mai di questo termine.

Quando, trent'anni piú tardi, Coulomb scoprí la legge d'interazione delle cariche elettriche, ipotizzò due tipi diversi di materia elettrica, ognuno formato da atomi indistruttibili, ma si spinse oltre usando la terminologia di Franklin e le definí cariche elettriche positive e negative. Ecco una splendida illustrazione del significato della terminologia: da una parte l'uso di Franklin del termine

'elettricità negativa' (non 'mancanza di', 'deficit', ecc.) poteva facilmente creare l'impressione che fosse implicata un'entità differente e questo avrebbe potuto influenzare Coulomb. D'altra parte per Coulomb vi sono realmente due entità diverse, cosicché la spiegazione frankliniana del processo di neutralizzazione non regge piú. Però l'impiego delle parole 'positivo' e 'negativo' dissimulò il problema e l'analogia con i numeri portò alla convinzione che fosse semplicemente naturale l'annullamento di quantità uguali di cariche positive e negative.

Coulomb stesso non si lasciò fuorviare dalla sua notazione, e cosí pure, dopo di lui, Faraday e Maxwell. Tutti si trovarono di fronte al problema della neutralizzazione. Come possono due sostanze diverse neutralizzarsi senza violare la legge di conservazione della materia, che a quell'epoca era considerata un principio al di là di qualsiasi dubbio? Questo problema era talmente serio che a lungo la teoria dualistica di Coulomb non conquistò una preferenza assoluta rispetto alla teoria monistica. Sperimentalmente era difficile distinguere tra le due teorie. Sia Faraday sia Maxwell furono molto chiari in proposito e, pur usando il modello dualistico, affermarono esplicitamente di non riuscire a confutare la teoria monistica.

Lasciando aperto il problema della conciliazione tra neutralizzazione e conservazione della materia, l'impressionante lavoro sperimentale di Faraday portò a una differente formulazione della legge di conservazione della carica: le cariche elettriche possono soltanto essere formate in coppie di cariche positive e negative di uguale forza. Ciò significa che la carica netta di un sistema isolato rimane costante, poiché ogni carica creata nel sistema (in conseguenza di attrito o di qualsiasi altro processo) sarà accompagnata da una carica uguale di segno opposto, in modo tale che la carica totale netta sia invariante.

Sarebbe opportuno aggiungere a questo punto alcuni commenti sulla legge:

a) Questa legge, pur formulata in termini dualistici, non esclude la possibilità dell'esistenza di un unico tipo di elettricità. Se le cose stanno cosí, la legge s'identifica con quella di conservazione della materia, poiché in un sistema isolato la materia non può essere creata ed è possibile soltanto una ridistribuzione della supposta materia elettrica. L'eccesso in una parte del sistema implica deficienza in un'altra parte, ovvero si formano cariche positive e negative uguali.

b) Come indicato in precedenza, alla legge di conservazione della carica si giunse per analogia con la conservazione della materia. Questa analogia non vale però per l'attuale formulazione. La legge di conservazione della materia negava la creazione e distruzione della materia, mentre la nostra legge non esclude alcuno dei due processi. Essa è soltanto una condizione sui *rapporti* delle differenti cariche create. Per di piú, la legge di conservazione della carica di fatto sopravvisse a quella di conservazione della materia. Oggi la legge di conservazione della materia non è piú ritenuta esatta, mentre quella di conservazione della carica è ancora considerata una delle leggi fisiche fondamentali.

c) La legge in questa formulazione non dice nulla sulla natura dell'elettricità, non si affida a una teoria dell'elettricità come materia, come materia atomistica, come proprietà della materia, ecc. Questo è uno dei motivi per cui questa legge sopravvisse a tutte le rivoluzioni che hanno subito le teorie della materia.

d) Concettualmente la legge non rappresenta un miglioramento rispetto alla precedente versione di Franklin. Forse è vero anche l'opposto: la legge sembra piú un assioma, naturalmente un assioma ben confermato, dunque un corollario di una teoria. Ciò non ne diminuisce in alcun modo il valore come legge scientifica – ogni teoria prende l'avvio da assiomi – ma dimostra ancora una volta che le leggi fondamentali non sono necessariamente fondate su concetti ben definiti; non vi è alcun preciso concetto di carica prima di quella legge, anzi è la legge stessa, questo assioma, uno dei fattori principali dell'emergere del concetto di carica.

e) Piú di tutte le altre leggi di conservazione classiche, questa è il risultato di un enorme lavoro sperimentale. Essa non venne giustificata metafisicamente come la legge di conservazione del moto di Descartes, né fa parte della nostra concezione intuitiva del mondo come si potrebbe considerare la legge di conservazione della materia. È uno dei casi, forse una delle eccezioni che confermano la regola, in cui ha veramente funzionato il metodo ipotetico-deduttivo.

2.5. Conservazione dell'energia.

La legge di conservazione dell'energia fu l'ultima delle leggi di conservazione classiche a essere scoperta. La sua storia fu lunga e complessa, e il concetto implicato – l'energia – fu il piú confuso tra tutti i concetti fondamentali classici. Ciò nonostante, o forse proprio per questo, la legge di conservazione dell'energia ebbe un impatto straordinario e penetrò profondamente nelle idee popolari, andando perciò soggetta a molti impieghi superficiali e imprecisi. Tuttavia, la fortuna di cui godette sul finire dell'Ottocento non la salvò da modificazioni sostanziali poco dopo il volgere del secolo.

La formulazione della legge di conservazione dell'energia è analoga a quella delle altre leggi di conservazione ed è cosí espressa: l'energia totale in un sistema chiuso è costante per qualsiasi modificazione subita dal sistema. Si deve ricordare che soltanto un sistema ideale è veramente un «sistema chiuso» rispetto all'energia. Mentre è possibile proteggere con notevole efficacia un sistema contro le perdite di carica elettrica o di materia, nessun isolamento impedirà dei trasferimenti di energia, poiché questi possono verificarsi anche nel vuoto.

Il principio di conservazione dell'energia afferma piú della semplice indistruttibilità di un'entità. Vi sono molte forme di energia e il principio stabilisce che esse sono equivalenti e trasformabili l'una nell'altra, pur restando costante la loro somma totale. La scoperta della legge richiese perciò lo studio di una moltitudine di fenomeni e la percezione dell'esistenza di un denominatore comune a tutti. Questo è uno degli elementi che complicarono la scoperta della legge. Inoltre, il fatto che fossero coinvolte forme diverse di energia andò ad aggiungersi alla confusione già esistente nei concetti e nella terminologia. La letteratura del XVIII e XIX secolo conosce un gran numero di presunte entità conservate: 'vis viva', 'funzione di forza', 'forza', 'potenziale', 'lavoro', 'tensione', ecc. Onde sostituire questi vaghi principî con un concetto ben definito di energia era necessaria una formulazione inequivocabile della legge in termini matematici.

Una ricostruzione razionale e schematica della storia della legge di conservazione dell'energia si potrebbe rappresentare cosí: Ipotesi *a*: l'energia si conserva; confutazione *a*: l'energia va perduta in un processo di tipo X_1. Ipotesi *b*: nel processo di tipo X_1 l'energia non è perduta ma semplicemente trasformata da energia di tipo A in energia di tipo B. La somma totale dell'energia A e dell'energia B si conserva; confutazione *b*: l'energia è perduta in un processo di tipo X_2... Ecc. Ma questa è una ricostruzione possibile soltanto col senno di poi. Il fatto che ogni ipotesi fosse un gradino verso la legge generale di conservazione dell'energia non avrebbe potuto essere compreso prima della pratica scoperta della legge. Anche il concetto di energia, liberamente usato nella ricostruzione, entrò in uso soltanto *dopo* la divulgazione della legge e sino ad allora non si manifestò alcun concetto isolato comune a tutte le varie ipotesi.

Inoltre, non è possibile dare una definizione del concetto di energia indipendentemente dalla legge di conservazione. Ogni stadio dello schema rende possibile la definizione di un concetto – un tipo di energia. L'impiego dello stesso termine 'energia' per un'altra quantità è giustificato soltanto dalla legge che asserisce che la nuova quantità è semplicemente una versione trasformata della precedente. Piú precisamente, attribuire all'energia il significato di capacità di compiere lavoro (una definizione piuttosto diffusa) ha senso soltanto se questa «capacità» può essere attualizzata e se sussiste una relazione matematica invariante tra la «capacità» e il lavoro meccanico prodotto. Ma questo è esattamente ciò che rivendica la legge di conservazione. L'unica possibile definizione è quindi una definizione basata sulla legge.

La legge di conservazione dell'energia, come pure la legge di conservazione del momento si fondavano sulla persuasione che nel moto qualcosa si conservasse. Com'è stato detto, la quantità di moto conservata di Descartes precorreva il momento mv di Newton. D'altra parte Leibniz postulò la conservazione di una quantità proporzionale alla quantità di materia e al quadrato della velocità, ossia mv^2 e la definí *vis viva* 'forza viva'. Egli notò anche una connessione tra quella quantità e l'altezza raggiunta da un corpo scagliato verso l'alto: entrambe sono «effetti» della medesima forza. È questa forza, pur generale e indefinita, che Leibniz supponeva venisse conservata. La legge meccanica di conservazione della *vis viva* è soltanto un caso di una legge metafisica piú generale. Questo è chiaro se si tiene presente l'argomentazione metafisica generale con cui Leibniz sostiene la sua legge. Per Leibniz vi sono due principî esplicativi: il principio d'identità e il principio di ragion sufficiente. Quest'ultimo, che s'identifica in realtà con la legge di causalità, esige l'uguaglianza tra causa e effetto. Tale principio viene usato da Leibniz per dimostrare la conservazione della *vis viva*. La *vis viva* non può aumentare (ciò significherebbe che l'effetto è maggiore della causa); d'altra parte, se decrescesse soltanto, alla fine dovrebbe svanire del tutto, il che è contrario all'ordine dell'universo e alla perfezione di Dio. Ne consegue che la *vis viva* si deve conservare. La conservazione della forza come manifestazione della potenza e perfezione di Dio è uno dei punti di discussione principali nella controversia tra Leibniz e Newton, quale è riflessa nella corrispondenza Leibniz-Clarke. Per entrambi la «forza» è una delle entità fondamentali, ma an-

cora una volta la terminologia è ingannevole: la forza di Newton è molto diversa da quella di Leibniz e la famosa corrispondenza non è in realtà una discussione sulla stessa quantità fisica, ma piuttosto un confronto tra due dottrine scientifiche diverse. Come si è già osservato, le leggi di conservazione scaturiscono facilmente dalla meccanica di Newton, ma egli non le derivò mai come tali.

La quantità chiamata da Leibniz *vis viva* – a metà del cui valore si dà il nome di energia cinetica – si conserva soltanto nelle collisioni elastiche. Questo principio qualificato era già stato scoperto da Huygens nel 1699. Leibniz, per il quale il principio era assoluto, dovette rendere ragione della perdita di *vis viva* nelle collisioni inelastiche. La sua spiegazione fu che la forza veniva impartita agli atomi costituenti il corpo, cosicché ciò che era perduto dal corpo veniva guadagnato dalle sue parti interne. Tale spiegazione non ebbe vasta risonanza e la perdita di *vis viva* nelle collisioni inelastiche continuava a costituire un problema nel 1847, anno della scoperta della legge di conservazione dell'energia.

Nel corso del XVIII secolo la scienza della meccanica subí un'importante trasformazione dalla pura formulazione vettoriale di Newton alla chiara formulazione scalare di Lagrange. Al posto della forza newtoniana **F** nella teoria di Lagrange si trova una «funzione di forza» (oggi energia potenziale) dalla quale deriva la forza:

$$\mathbf{F} = -\nabla V = -\left(\frac{\partial V}{\partial x} + \frac{\partial V}{\partial y} + \frac{\partial V}{\partial z}\right).$$

Anziché il momento vettoriale, si ha l'energia cinetica $T = mv^2/2$ e l'equazione di Lagrange è l'equazione del moto dello scalare combinato $L = T - V$. La legge di conservazione dell'energia meccanica è effettivamente incorporata nel formalismo lagrangiano. L'energia meccanica del sistema è $T - V$ e si può facilmente provare che questa quantità si conserva. L'espressione ammette però un'importante restrizione: è valida solo per sistemi conservativi, ossia sistemi in cui il potenziale non dipende dalle velocità, e per sistemi nei quali i vincoli sono indipendenti dal tempo. Poiché le forze d'attrito dipendono dalle velocità, l'energia meccanica, secondo Lagrange, può essere perduta in qualsiasi sistema soggetto ad attrito; in altre parole l'energia si conserverà esclusivamente in un sistema ideale.

Illustriamo la connessione tra energia potenziale ed energia cinetica con un semplice esempio: se una forza costante **F** agisce su un corpo lungo una distanza **x**, essa produce un lavoro $W = \mathbf{F} \cdot \mathbf{x}$. Il corpo verrà accelerato in base alla seconda legge di Newton $\mathbf{F} = m \cdot \mathbf{a}$ (m è la massa, **a** l'accelerazione). La distanza **x** è data da $\mathbf{a}t^2/2$ (t per secondi) e quindi il lavoro produce energia cinetica secondo la formula $W = ma^2t^2/2 = mv^2/2$. Supponiamo ora che la forza sia la forza gravitazionale e che l'accelerazione sia l'accelerazione di gravità. Si può allora affermare che un corpo posto ad altezza x da terra ha un'energia potenziale $m \cdot a \cdot x$ che si trasformerà in energia cinetica $mv^2/2$ nel momento in cui raggiungerà il suolo. D'altra parte, se il corpo viene scagliato in alto con energia cinetica iniziale $mv^2/2$, proseguirà in quella direzione soltanto per un tratto x, poiché a quel momento tutta la sua energia cinetica si è convertita in energia potenziale. Una serie

di trasformazioni successive da energia cinetica in energia potenziale costituisce il moto di un pendolo oscillante. Si noti che il concetto di energia potenziale è in pratica collegato a un sistema e non a un singolo oggetto. L'oggetto di cui sopra possiede energia potenziale in virtú della forza che su di esso esercita la Terra. L'energia potenziale $m \cdot a \cdot x$ è in realtà l'energia del sistema oggetto-Terra separato dalla distanza x. L'energia cinetica fu anche misurata rispetto alla Terra e non, ad esempio, rispetto alle stelle fisse. Inoltre si può scegliere arbitrariamente il punto zero dell'energia potenziale.

Un altro problema molto discusso nel XVIII secolo fu la possibilità o meno del «moto perpetuo».

La costruzione di una macchina in grado di funzionare continuamente senza alimentazione di energia fu una sfida che molti cercarono di raccogliere, sia tra gli scienziati sia tra gli artigiani. Nel 1775 l'Académie des sciences francese votò una risoluzione per non esaminare ulteriori tentativi di costruzione di una simile macchina. Se l'Académie avesse riconosciuto una legge di conservazione dell'energia, la decisione senz'altro sarebbe stata basata su di essa, ma tra le ragioni pubblicate dall'Académie non figura alcun principio di questo genere. Vengono forniti due argomenti: *a*) l'insuccesso totale di tutti i tentativi che erano stati esaminati, ossia un argomento basato su una generalizzazione induttiva; *b*) il principio scolastico sopra citato secondo cui la causa deve uguagliare l'effetto. Questo principio fu adottato per significare che se la macchina deve lavorare sulla scorta della prima forza motrice impressale, essa può produrre soltanto un effetto infinitamente piccolo ad ogni istante, sarebbe cioè una macchina completamente inutile. Questo argomento si applica anche in assenza di attrito e resistenza, e quindi a maggior ragione se si prendono in considerazione tali forze. L'Académie è ansiosa di sottolineare che non vi è contraddizione tra l'impossibilità di una macchina in moto perpetuo e il principio d'inerzia, secondo cui nello spazio vuoto un corpo conserverà il suo moto uniforme, che però per sua stessa definizione non può produrre lavoro. A noi la connessione tra questa risoluzione e la legge di conservazione dell'energia sembra talmente ovvia che spesso si è affermato che la legge sia derivata proprio dall'impossibilità del moto perpetuo. Tuttavia, se si tiene conto del fatto che prima dell'enunciazione della legge passarono altri settant'anni, bisogna concludere che al massimo fu uno dei fattori che contribuirono alla creazione del clima scientifico in cui operarono gli scopritori della legge nel XIX secolo.

Riassumendo lo sviluppo del principio di conservazione sino al 1788 (anno in cui comparve la *Mécanique analytique* di Lagrange): fu riconosciuta l'impossibilità del moto perpetuo e venne stabilita la legge di conservazione dell'energia meccanica per sistemi non implicanti attrito. Ma cosa accade quando sono presenti le forze d'attrito? L'energia va in tal caso perduta? Nel 1798 il conte Rumford lesse alla Royal Society la sua famosa relazione sugli esperimenti compiuti per provare che il calore si ottiene per attrito. Avendo osservato che un fucile di ottone si scalda notevolmente mentre viene alesato, egli si accinse a scoprire l'origine di quel calore. Il problema che interessava Rumford era la natura del calore. La teoria del calore prevalente a quell'epoca era quella calorica, secondo la quale

il calore è un fluido materiale. Questa teoria venne definita da Rumford teoria moderna, in contrapposizione alla piú antica teoria che considerava il calore come movimento ed era sostenuta da Bacone, Locke, Hooke, Newton e Leibniz. Alcuni dei maggiori scienziati del XVIII secolo erano esitanti sulla natura del calore. Cosí anche Lavoisier, incline ad accettare la teoria calorica; pubblicò un articolo in collaborazione con Laplace in cui il calore è definito come *vis viva* delle molecole. Analogamente, Davy, di solito ritenuto, insieme a Rumford, uno degli scopritori della teoria del calore come movimento, nutriva seri dubbi e pensava che le prove fossero troppo scarse per decidere tra le due teorie. Rumford dimostrò che il calore non poteva essere un fluido materiale poiché un simile fluido avrebbe dovuto rispettare la legge di conservazione della materia, sicché la quantità di calore prodotto in un dato sistema materiale doveva essere limitata. I suoi esperimenti provarono tuttavia che mediante attrito è possibile produrre calore a piacere, risultato in accordo con la teoria moderna. Il lavoro di Rumford, cosí significativo per la scoperta della legge di conservazione dell'energia, era quindi di fatto basato sulla legge di conservazione della materia. Rumford non collegò le sue osservazioni con la legge già esistente della conservazione dell'energia meccanica. Il legame tra calore e lavoro meccanico fu stabilito verso la metà del XIX secolo da parecchi scienziati impegnati in programmi di ricerca diversi. Joule studiò il funzionamento della cella elettrolitica e della macchina elettromagnetica. Le sue indagini lo portarono a rendersi conto dell'esistenza di una stretta connessione tra il lavoro meccanico realizzabile da una macchina, la corrente elettrica nella macchina e il calore prodotto da quella corrente. Il suo lavoro sperimentale, di un rigore fuori del comune, culminò nel calcolo dell'equivalente meccanico del calore. In unità CGS (e con un miglioramento molto lieve rispetto ai risultati di Joule) questo equivalente vale $4{,}187 \times 10^7$ erg per una caloria (la caloria è la quantità di calore necessaria per elevare di un grado centigrado la temperatura di un grammo d'acqua a pressione atmosferica normale). Joule dimostrò anche che la quantità di calore prodotta da una corrente elettrica nella macchina elettromagnetica era identica se il lavoro usato per produrre la corrente era stato trasformato *direttamente* in calore per attrito. Gli esperimenti di Joule risolsero la disputa sulla teoria del calore e aprirono la strada alla moderna termodinamica. La conclusione cui giunse Joule sulla base del suo lavoro sperimentale fu che la forza viva non viene annullata dall'attrito: ogniqualvolta essa viene apparentemente distrutta, si produce un equivalente, che a sua volta può essere di nuovo convertito in forza viva. Questa conclusione non era convalidata soltanto da esperimenti. Joule era profondamente convinto che la distruzione della forza viva fosse impensabile, al punto che riteneva possibile pervenire alla stessa verità mediante un ragionamento a priori, basato sulla conoscenza della potenza e saggezza di Dio.

Un altro approccio al problema venne da un interesse d'ordine fisiologico, rappresentato dalla relazione tra il calore degli organismi viventi, il lavoro meccanico che essi producono e la combustione del cibo che consumano. Julius Robert von Mayer, un medico, osservò che nelle zone tropicali si aveva una variazione del colore del sangue venoso; egli attribuí tale cambiamento a un eccesso di ossigeno nel sangue, conseguente alla ridotta combustione di alimenti. Ciò lo indusse a una serie di esperimenti, nei quali tentò di misurare l'equivalente meccanico del calore, pubblicati nel 1842 e nel 1845. I suoi risultati erano un po' diversi da quelli di Joule, ma ciò che interessa è la conclusione: le forze della natura sono indistruttibili, ogni perdita è solo apparente. Si osservi che il concetto fondamentale di Mayer era la «forza» (*Kraft*), mentre per Joule predominavano il «lavoro» e la *vis viva*. Anche Mayer aveva un principio-guida metafisico: le forze sono cause, argomenta Mayer, e le cause debbono uguagliare gli effetti; quindi le forze sono indistruttibili.

Ciò che ancora mancava, nonostante la crescente persuasione di qualcosa d'indistruttibile, e malgrado i precisi dati sui processi di trasformazione del lavoro, era la prova matematica che in tutti i processi fisici esiste una quantità avente sempre la stessa dimensione e conservante sempre la propria grandezza totale. In breve ciò che mancava era la legge di conservazione dell'energia. Fu Helmholtz a scoprire questa legge nel 1847 e a pubblicarla (privatamente, poiché fu rifiutata da una rivista) col titolo *Sulla conservazione della forza* (*Über die Erhaltung der Kraft*).

L'articolo di Helmholtz si apre con una discussione sulla legge di conservazione in meccanica. Un sistema possiede *vis viva* (energia cinetica) in virtú del suo moto; è dotato di tensione (energia potenziale) per via della sua struttura e delle forze interne. La somma totale della «tensione» e della *vis viva* si conserva. Come corollario, Helmholtz dimostra l'impossibilità del moto perpetuo. La parte originale del lavoro è quella che si occupa di altre manifestazioni della *Kraft*. Nelle collisioni inelastiche e nei sistemi ove agiscono forze d'attrito la perdita di *vis viva* è associata alla produzione di tensione, una delle cui forme è il calore. Il calore è la *vis viva* delle parti interne della materia. Helmholtz s'interessa anche dell'equivalente della *Kraft* per processi elettrici, magnetici ed elettromagnetici. Egli sfiora il problema del bilancio energetico degli organismi viventi e, pur non giungendo ad alcuna conclusione definita, avanza l'ipotesi che gli animali consumino tensione chimica e producano calore e lavoro meccanico in base alla legge di conservazione. Alcuni semplici esempi aiuteranno a spiegare come sia definita l'energia potenziale in campi diversi da quello meccanico: l'energia elettrostatica di un corpo caricato è il lavoro compiuto per caricarlo (si deve spendere lavoro per caricarlo in quanto ogni carica sul corpo respingerà ulteriori cariche in base alla legge di Coulomb). Se il conduttore assume una carica Q e il suo potenziale elettrico finale è V, il lavoro compiuto è espresso da $QV/2$.

Se la resistenza di un conduttore è R, si debbono spendere I^2R unità di energia al secondo per mantenere una corrente I. La corrente a sua volta produrrà la stessa quantità di calore: I^2R unità di calore al secondo.

Ogni settore della fisica ha la propria energia potenziale in accordo con il principio generale. Molle, fluidi elastici e gas, onde sonore, onde luminose, tutte hanno proprie specifiche energie potenziali ed è impossibile soffermarsi separatamente su ogni definizione. È importante sceverare questi concetti di energia, impiegati nella fisica e soggetti a formulazione matematica, da altri usi che allo stadio attuale sembrano piuttosto metaforici, come ad esempio «energia men-

tale». Ciò non implica affatto che gli usi metaforici non possano essere fruttuosi nella loro sfera d'azione.

Due problemi continuarono a tenere occupati i fisici dopo Helmholtz e Joule: 1) la differenziazione della *Kraft* helmholtziana da altre presunte quantità conservate quali la forza di Faraday; 2) la riconciliazione tra la legge di conservazione dell'energia e il principio di Carnot (scoperto da Sadi Carnot), che pone un limite al rendimento di una macchina termica e quindi suggerisce che il calore non sia completamente trasformabile in lavoro meccanico. Il primo problema fu trattato da William Rankine in un articolo dal titolo *On the General Law of Conservation of Energy*, pubblicato nel 1853, dove introdusse il termine 'energia' al posto di quello che Helmholtz chiamava *Kraft* e illustrò come i risultati di Joule derivino dal principio di Helmholtz. Il secondo problema fu risolto da Lord Kelvin nel suo articolo del 1851 *On the Dynamical Theory of Heat* e da Rudolf Clausius pressappoco nello stesso periodo. Si dimostrò che non vi era alcuna contraddizione tra i due principî: in ogni processo *attuale* l'energia viene senz'altro conservata, ma non tutti i processi teoricamente possibili, in base alla legge di conservazione dell'energia, si verificano realmente in natura (o meglio, hanno luogo abbastanza spesso da consentirne l'osservazione). Il principio di Carnot, o seconda legge della termodinamica, stabilisce quali processi si realizzino in pratica.

Dopo lunga lotta la legge di conservazione dell'energia infine ottenne un riconoscimento universale, anzi finí con l'essere considerata come il principio fondamentale della fisica e col dar vita a tutta una scuola che patrocinò la scienza dell'energetica. I seguaci della nuova scienza, tra i quali Ostwald, credevano che tutti i fenomeni fossero spiegabili come trasformazioni di un'entità fondamentale – l'energia – piú fondamentale della materia sia dal punto di vista ontologico sia da quello epistemologico. La successiva metamorfosi del concetto di energia è discussa nel § 3.

2.6. Leggi di conservazione e simmetrie nella fisica classica.

Dopo Lagrange, nel corso del XIX secolo, la meccanica fu riformulata varie volte con crescente eleganza matematica. I maggiori contributi vennero da Hamilton e da Jacobi. Queste nuove esposizioni non modificarono il contenuto della meccanica ma non furono prive di significato fisico, in quanto resero possibili nuove intuizioni e piú ampie applicazioni. Ad esempio ci si accorse che le leggi di conservazione dell'energia e del momento sono derivabili da considerazioni molto semplici sulle proprietà di simmetria dei sistemi fisici e dello spazio e del tempo.

La funzione essenziale nel formalismo di Hamilton, piú tardi definita hamiltoniana, è una funzione che in determinate condizioni rappresenta l'energia totale del sistema e contiene tutte le informazioni sulla produzione del moto del sistema nel tempo. La sua definizione matematica è:

$$H=\sum_i \dot{q}_i p_i - L,$$

ove q_i è la coordinata $i^{-\text{esima}}$ del sistema, $\dot{q}_i$ è la sua derivata rispetto al tempo e infine p_i è il momento coniugato all'$i^{-\text{esima}}$ coordinata. L è la lagrangiana nel senso definito sopra. Si può dimostrare che, se non è una funzione esplicita del tempo, la hamiltoniana è una costante di moto o in altri termini è *invariante* lungo tutto lo sviluppo temporale del sistema. Si può inoltre rilevare che se la hamiltoniana non è una funzione esplicita di altra coordinata, allora il momento coniugato alla $i^{-\text{esima}}$ coordinata si conserva, o è una costante di moto. La fisica del sistema è contenuta nella hamiltoniana; tutto il resto è una derivazione matematica. È la fisica del sistema a determinare quali coordinate entrino esplicitamente nella hamiltoniana. Se il sistema è simmetrico per una certa trasformazione, tale simmetria deve trovare la propria espressione nella hamiltoniana. Ad esempio una traslazione di un sistema libero nello spazio non può modificare la fisica della situazione, poiché è equivalente alla variazione dell'origine del sistema di coordinate mentre il sistema resta dove si trova. L'affermazione che un tale spostamento dell'origine non modifica la situazione fisica si basa sull'assunto che lo spazio è omogeneo. Secondo queste considerazioni la coordinata del baricentro del sistema non può comparire esplicitamente nella hamiltoniana e di conseguenza il momento totale del sistema si deve conservare. Analogamente, se un sistema libero è simmetrico rispetto a un certo asse, certamente una rotazione attorno a quell'asse non varierà la situazione fisica: qui si suppone che lo spazio sia *isotropico*, ovvero privo di direzioni preferenziali. In tal caso la hamiltoniana non sarà una funzione esplicita dell'angolo di rotazione attorno all'asse e, di conseguenza, il momento coniugato, cioè il momento angolare attorno all'asse, verrà conservato. È da notare quindi che la legge di conservazione del momento è una conseguenza dell'omogeneità dello spazio. Per analogia si può affermare che la legge di conservazione dell'energia è un'espressione di una omogeneità fondamentale nello sviluppo del sistema nel tempo. Questa connessione tra leggi di conservazione e simmetrie è divenuta ancora piú importante nella fisica moderna, come si vedrà in seguito.

3. *Leggi di conservazione e invarianza nella fisica contemporanea.*

Nel paragrafo precedente si è descritto il modo in cui la meccanica classica, dopo un lungo processo di sviluppo, è giunta a un formalismo nel quale erano incorporate le leggi di conservazione e le proprietà di invarianza, collegate a simmetrie. Questo formalismo fu il punto di partenza della fisica del XX secolo, cosicché le leggi di conservazione sono inserite nella teoria esattamente dal principio. Alcune delle leggi di conservazione classiche furono accettate, altre modificate o addirittura del tutto trasformate. Sono anche state scoperte e incorporate nel formalismo molte nuove leggi di conservazione.

La fisica contemporanea si fonda sulla teoria della relatività e sulla meccanica quantistica. Mentre lo sviluppo storico di queste teorie è discusso piú particolareggiatamente negli articoli loro dedicati, la discussione che segue si concentra sugli aspetti fisici dell'argomento. Dopo aver trattato (§ 3.1) delle leggi di con-

servazione nella teoria della relatività speciale (la teoria della relatività generale non è qui discussa) si discuteranno le leggi di conservazione nella meccanica quantistica, iniziando con alcune notazioni sulla vecchia teoria dei quanti (§ 3.2), per poi procedere a esaminare la meccanica quantistica contemporanea e la sua descrizione delle particelle elementari (§§ 3.3 sgg.).

3.1. Relatività e invarianza.

La teoria della relatività speciale si basa su due assunti: *a*) le leggi naturali sono identiche, ovvero hanno la stessa forma matematica, in tutti i sistemi in moto uniforme l'uno rispetto all'altro; *b*) la velocità della luce è costante indipendentemente dal moto dell'osservatore rispetto alla sorgente. La prima ipotesi era già uno dei presupposti della meccanica classica (piú spesso implicita che esplicita) ma, soltanto quando fu associata alla seconda, ad opera di Einstein, poterono nascere i nuovi risultati della teoria della relatività. Sulla base di questi due assunti la teoria della relatività costruisce una trasformazione di coordinate da un sistema mobile a un altro, definita trasformazione di Lorentz, e il postulato basilare della teoria afferma che le leggi di natura sono invarianti rispetto alla trasformazione di Lorentz. Il concetto di invarianza di una legge fisica potrebbe a prima vista apparire involuto. Il concetto di legge non significa di per sé che la legge è in generale vera e non varia da un sistema di riferimento a un altro? Un esame piú attento chiarisce però che il concetto di legge implica semplicemente un comportamento simile in condizioni analoghe, e per ogni particolare legge si deve specificare quali siano quelle condizioni analoghe. La richiesta che in tutti i sistemi in moto uniforme le leggi debbano essere identiche è perciò ben lontana dall'essere banale. Le leggi fisiche non hanno affatto la stessa forma matematica in *tutti* i sistemi di riferimento; in sistemi accelerati o ruotati l'uno rispetto all'altro le leggi della meccanica presentano una forma diversa (questo non è vero secondo la teoria della relatività generale).

La teoria della relatività opera in uno spazio vettoriale tetradimensionale noto come spazio di Minkowski. Le prime tre componenti di un vettore di posizione sono le coordinate spaziali e la quarta è una coordinata immaginaria proporzionale al tempo, $X_y = ict$, ove t è il tempo, c la velocità della luce e $i = \sqrt{-1}$. Il principio d'invarianza per la trasformazione di Lorentz significa che le leggi fisiche sono invarianti rispetto a rotazioni nello spazio di Minkowski.

Come è stato detto, la trasformazione di Lorentz venne costruita sulla base dell'assioma che la velocità della luce sia costante. Come conseguenza di questa trasformazione alcune quantità *invarianti* nella meccanica classica si rivelano *varianti* da un sistema di riferimento a un altro. Ad esempio un intervallo di tempo è piú lungo se lo si misura in un sistema in movimento anziché in un sistema in quiete (naturalmente ognuno dei due sistemi può essere considerato fermo da un osservatore posto in esso) e la lunghezza di ogni oggetto situato lungo la direzione del movimento è minore in un sistema mobile.

Il principio d'invarianza rispetto alla trasformazione di Lorentz vuole che tutte le leggi classiche non invarianti per la trasformazione vengano modificate e ricevano una forma invariante. In altre parole, si dovrebbe considerare il principio d'invarianza come una meta-legge, ovvero una legge sulle leggi fisiche, una legge che stabilisce le condizioni cui deve ottemperare qualunque legge fisica.

Una delle leggi che deve essere cambiata per soddisfare i requisiti del principio d'invarianza è la seconda legge del moto di Newton, dalla quale discende le legge di conservazione del momento. Poiché questa legge si accorda all'esperienza nel campo delle velocità molto inferiori a quella della luce, essa dev'essere modificata in modo tale che la nuova legge coincida con la precedente nel caso limite di v molto piú piccolo di c. Ne consegue che, se definiamo il momento in modo diverso dalla meccanica classica, il nuovo momento obbedisce a una legge analoga a quella di Newton e si conserva nelle stesse condizioni del momento classico. Tale momento relativistico è definito dalla formula:

$$p_i = \frac{mv_i}{\sqrt{1-\beta^2}} \qquad \text{per } i = 1, 2, 3$$

ove p_i è la componente i-esima del momento; v_i la componente i-esima della velocità; m la massa; e $\beta = v/c$ (v = velocità; c = velocità della luce). È da notare che, quando v è molto minore di c, possiamo trascurare β^2; in tal caso il momento relativistico s'identifica col momento classico. Questa formulazione fornisce soltanto tre componenti del vettore momento. La quarta componente, necessaria per definire un vettore tetradimensionale relativistico, è iT/C, essendo $T = mc^2/\sqrt{1-\beta^2}$ l'energia cinetica relativistica dell'oggetto. Quest'energia cinetica relativistica non s'identifica con l'energia cinetica classica, neppure per velocità basse. Il suo valore limite è $mc^2 + (mv^2/2)$, ossia la somma tra l'energia cinetica classica e un altro fattore definito energia a riposo. Le quattro componenti cosí definite costituiscono il vettore momento tetradimensionale e seguono *una* legge di conservazione, la legge di conservazione del momento-energia. Quindi ogni volta che il momento si conserva anche l'energia cinetica dev'essere conservata. Nella meccanica classica questo principio non è valido, ad esempio in una collisione inelastica il momento si conserva ma non cosí l'energia cinetica, e si produce calore per equilibrare la perdita. Secondo la teoria della relatività comunque l'energia cinetica deve conservarsi in una collisione inelastica. Ora, poiché $mv^2/2$ diminuisce, mc^2 deve aumentare e, dal momento che c^2 è costante, soltanto un incremento della massa a riposo m può essere responsabile dell'aumento. Ciò significa che la massa di un sistema non è piú una costante di moto. Massa ed energia si fondono in una variabile dinamica, la massa-energia, ed è questa variabile unificata che obbedisce a una legge di conservazione. La legge classica di conservazione della materia viene perciò abbandonata dalla teoria della relatività. Inoltre, la legge di conservazione dell'energia e la legge di conservazione del momento, separate nella meccanica classica, sono incorporate dalla teoria della relatività in un'unica legge: la legge di conservazione del vettore tetradimensionale del momento energia.

Talvolta è utile introdurre una massa relativistica, diversa dalla massa a riposo ed espressa dalla relazione $m_{\text{relat}} = m/\sqrt{1-\beta^2}$. I motivi sono puramente

formali. È importante non confondere i due concetti, poiché la massa relativistica di un sistema cresce con la velocità, mentre la massa a riposo è indipendente da questa. L'unificazione dei concetti di massa ed energia ha conseguenze di vasta portata. Il suo significato fisico è che la massa è trasformabile in energia e viceversa. La relazione tra variazione energetica e incremento della massa a riposo è data dalla famosa formula di Einstein: $E=\Delta mc^2$. Nella vita quotidiana le energie sono troppo basse perché i cambiamenti di massa siano significativi. Ad esempio nella combustione di una tonnellata di carbone la perdita di massa è inferiore a tre milligrammi! Tuttavia nel campo delle energie nucleari l'effetto delle trasformazioni di massa-energia può essere enorme. In linea generale l'energia che tiene unito un nucleo è molto elevata, ma succede che non tutti i nuclei siano ugualmente stabili. I piú stabili sono i nuclei con peso atomico medio, nei quali la massa è inferiore alla somma delle masse delle particelle che li costituiscono. Vi sono due modi di utilizzare questo fenomeno. Il primo è la fissione di un nucleo pesante, instabile, e il secondo è la fusione di nuclei molto leggeri in un nucleo piú stabile e pesante. Entrambi i procedimenti portano a una perdita di massa che si trasforma in energia. L'enorme quantità di energia liberata nella fissione e fusione dei nuclei si può impiegare, controllandola, per sostituire altre fonti energetiche. Sfortunatamente, l'applicazione principale sinora compiuta di questa energia è stata la fabbricazione della bomba atomica, che si basa sulla fissione del nucleo dell'uranio, e della bomba all'idrogeno, nella quale isotopi dell'idrogeno sono fusi in nuclei piú pesanti.

3.2. Le leggi di conservazione nella vecchia teoria dei quanti.

Una delle novità concettuali piú importanti della meccanica quantistica fu la descrizione della radiazione (luce, raggi X, ecc.) in termini di particelle, denominate quanti, analoghe sotto molti aspetti ai corpuscoli di materia. Questo nuovo modello richiese una revisione delle definizioni di energia e momento. Dal punto di vista delle leggi di conservazione, due effetti sperimentali svolsero un ruolo importante in questa rielaborazione:

a) L'effetto fotoelettrico spiegato da Einstein nel 1905: si è osservato che un metallo illuminato emette elettroni la cui energia cinetica dipende dalla frequenza della radiazione e non dalla sua intensità, come nella teoria ondulatoria classica, dove l'energia è proporzionale all'intensità dell'onda. Pareva che l'energia degli elettroni non dipendesse dall'energia della radiazione, il che era contrario alla legge di conservazione dell'energia. Einstein suppose che la radiazione fosse assorbita dal metallo in quantità discrete, con energia pari a $h\nu$ ove h è la costante di Planck e ν è la frequenza della radiazione. Ognuna di queste quantità è definita fotone. L'intensità della luce determina esclusivamente il numero di elettroni emessi e non la loro energia. La legge di conservazione dell'energia è rigorosamente mantenuta nell'intero processo.

b) L'effetto Compton scoperto da Compton nel 1923: quando raggi X o raggi γ sono diffusi da elettroni, si verifica un cambiamento nella loro lunghezza d'onda. Un modello ondulatorio per la radiazione non poteva dare una spiegazione del cambiamento. Compton suggerí quindi un modello corpuscolare, analogo a quello di Einstein per la luce. Le particelle di radiazione interagiscono con le particelle di materia in base alle leggi di conservazione del momento e dell'energia. I quanti di radiazione, venendo a collisione con gli elettroni, perdono un po' della loro energia e momento, cosicché la radiazione diffusa ha una lunghezza d'onda maggiore e gli elettroni acquistano energia e momento, rinculando di conseguenza. Questa teoria in pratica si discostava completamente dalla teoria ondulatoria della radiazione – passo che sembrò troppo rivoluzionario a Bohr, Kramers e Slater, i quali nel 1924 tentarono un approccio diverso. Essi cercarono di salvare la teoria ondulatoria, abbandonando le leggi di conservazione dell'energia e del momento per eventi individuali e interpretandole come leggi statistiche valide soltanto quando si prenda in esame un gran numero di eventi. Questa congettura fu confutata un anno dopo da Bothe e Geiger, i quali dimostrarono che le leggi di conservazione si mantengono per ogni singola collisione tra fotone ed elettrone.

3.3. Leggi di conservazione nella meccanica quantistica.

Negli anni 1925-26 vari nuovi formalismi equivalenti furono messi a punto per sostituire la meccanica classica e superare le difficoltà della vecchia teoria dei quanti. Nella nuova meccanica quantistica le variabili dinamiche quali il momento, l'energia, ecc. sono rappresentate da operatori che seguono regole diverse da quelle rispettate dalle variabili classiche. Il formalismo della meccanica quantistica per molti versi è analogo a quello hamiltoniano, di cui si è già parlato. Come nel formalismo hamiltoniano, la hamiltoniana del sistema, che in questo caso è un operatore, contiene le informazioni fisiche relative al sistema. In certi casi la hamiltoniana rappresenta l'energia totale del sistema ed ha forma analoga alla hamiltoniana classica; in molti altri casi non si può ancora calcolare la forma esplicita della hamiltoniana, il che costituisce uno dei gravi inconvenienti della meccanica quantistica. Se la hamiltoniana è nota, esiste un modo semplice per determinare i valori di probabilità degli operatori che rappresentano le variabili dinamiche. (È risaputo che la meccanica quantistica è una teoria statistica e quindi consente soltanto il calcolo di valori di probabilità).

Il formalismo matematico piú opportuno per trattare gli operatori della meccanica quantistica è l'algebra delle matrici. Una matrice di dimensione $m \cdot n$ è un insieme di $m \cdot n$ numeri disposti in m file e n colonne, e può essere affrontata come singola entità matematica soggetta a determinate regole di addizione e moltiplicazione. Il formalismo delle matrici della meccanica sarà discusso altrove; ai fini del presente articolo è importante solo un aspetto del formalismo: la differenza tra le variabili della meccanica quantistica e le variabili classiche trova la sua espressione nel carattere non commutativo dell'algebra delle matrici, e cioè nel fatto che in linea generale, per due matrici A e B, $A \cdot B$ è diverso da $B \cdot A$ o, in altre parole, la relazione di commutazione tra A e B ($A \cdot B - B \cdot A$) è in generale diversa dalla matrice zero.

C'è una relazione fondamentale tra il valore di probabilità di un osservabile

e la hamiltoniana, la quale afferma che la velocità di variazione nel tempo (cioè la derivata rispetto al tempo) del valore di probabilità di un'osservabile che non sia una funzione esplicita del tempo è proporzionale alla relazione di commutazione dell'osservabile con la hamiltoniana. Ne segue che, se l'osservabile commuta con la hamiltoniana, il suo valore di probabilità è una costante di moto. Per specificare le entità dinamiche conservate o le costanti di moto si debbono trovare gli operatori che commutano con la hamiltoniana. Si può dimostrare che un operatore di questo tipo dev'essere in grado di indurre sulla hamiltoniana una trasformazione che la lasci invariante. Quindi costanza di un operatore significa che la hamiltoniana è invariante rispetto alla trasformazione indotta da quell'operatore. Ad esempio la conservazione del momento lineare si può interpretare come un risultato dell'invarianza della hamiltoniana per una traslazione. Come si vedrà esistono trasformazioni rispetto alle quali la hamiltoniana è invariante, che non hanno una corrispondenza classica e che dànno luogo a nuove leggi di conservazione non presenti nella meccanica classica. Le leggi di conservazione del momento e dell'energia inoltre non sono del tutto analoghe alle leggi classiche, in quanto in molti casi la meccanica quantistica consente soltanto un gruppo discreto di valori possibili per variabili dinamiche che, nella meccanica classica, avevano uno spettro di valori continuo. Per di piú, nella meccanica classica lo stato di un sistema meccanico può essere completamente individuato esprimendo tutte le coordinate e tutti i momenti (e quindi anche tutti i momenti angolari). Nella meccanica quantistica, gli operatori che non commutano tra loro non possono essere specificati contemporaneamente, cosicché vi sono casi in cui il momento (o il momento angolare) del sistema non è ben definito, ovvero – nel linguaggio della meccanica quantistica – non è un buon numero quantico. In tali casi *non ha senso* annoverare queste variabili dinamiche tra quelle conservate.

3.4. Spin e statistica.

La meccanica quantistica ebbe subito molto successo nello spiegare gli spettri atomici; tuttavia si era notato che le variabili classiche (in questo caso l'energia e il momento angolare) non sono sufficienti per rendere ragione della suddivisione delle righe spettrali osservate. Questo fu uno dei motivi che condussero all'introduzione di una nuova variabile dinamica, senza corrispettivo classico, definita *spin*, dapprima per l'elettrone e piú tardi anche per altre particelle. Le regole algebriche cui ottempera lo spin sono quelle valide per il momento angolare orbitale, e quindi è possibile considerare lo spin come un momento angolare intrinseco. La situazione è abbastanza simile a quella della Terra nella sua rotazione attorno al Sole e contemporaneamente attorno al proprio asse. Questa analogia non dev'essere spinta troppo oltre, poiché nella meccanica classica entrambi i momenti risultano da una data situazione dinamica e non caratterizzano in alcun modo i corpi implicati. Lo spin è comunque una proprietà intrinseca di una particella indipendentemente dall'interazione in cui essa è coinvolta. In molti casi l'operatore dello spin commuta con la hamiltoniana ed è una costante di moto. Vi sono anche eventualità in cui il momento orbitale e il momento intrinseco non si conservano separatamente, ma la loro somma vettoriale commuta con la hamiltoniana ed è una costante di moto. Lo spin di una particella, al pari del momento angolare orbitale, può soltanto assumere valori discreti rispetto a qualsiasi asse prescelto; l'unica differenza sta nel fatto che mentre il momento angolare orbitale relativo a un dato asse può assumere soltanto valori interi multipli di $h/2\pi$, lo spin può avere valori interi o seminteri multipli della stessa costante. Ad esempio lo spin dell'elettrone misurato in unità $h2\pi$ vale $1/2$. Ciò significa che, scelta una direzione, lo spin dell'elettrone può soltanto ricevere i due valori $+(1/2)$ o $-(1/2)$ rispetto ad essa (al contempo altre componenti dello spin non saranno ben definite). Poiché assumerà sempre lo stesso insieme di valori, lo spin di una particella può fornire un mezzo per classificare le particelle. L'elettrone, il protone, il neutrone e molte altre particelle elementari hanno spin $1/2$. Il fotone ha spin 1 (ossia può acquistare i valori -1, 0, 1) e i mesoni (cfr. oltre) hanno spin 0. Si ha qui il primo esempio del piú profondo significato che le costanti di moto assumono nella meccanica quantistica: esse non soltanto forniscono un metodo opportuno per calcolare velocità, energie, ecc. come nella meccanica classica, ma definiscono anche la natura stessa degli oggetti fisici di cui ci si occupa.

Lo spin di una particella è strettamente legato al tipo di statistica cui sarà soggetta la particella quando fa parte di un sistema di molte particelle identiche. Il problema statistico è come dividere la particella del sistema tra i possibili stati quantici, ovvero come calcolare la probabilità che un certo numero di particelle occupi un dato stato quantico caratterizzato da un gruppo di numeri quantici (non si deve confondere questo problema statistico con l'interpretazione statistica della funzione d'onda).

Per illustrare il problema, cerchiamo di distribuire due particelle 1 e 2 tra due stati a e b. A prima vista sembra naturale distinguere quattro possibilità: 1 e 2 in a; 1 in a, 2 in b; 1 in b, 2 in a; 1 e 2 in b. Assegneremo quindi una probabilità $1/2$ al caso in cui 1 e 2 si trovano in stati diversi. Questo tipo di statistica è osservato dalle particelle classiche nella meccanica statistica classica e non si ritrova nella meccanica quantistica. Se aggiungiamo un postulato in base al quale la seconda possibilità è indistinguibile dalla terza (il che significa che le particelle non si possono contrassegnare, ossia mancano, per cosí dire, di individualità), esistono soltanto tre possibilità ed è necessario assegnare una probabilità $1/3$ al caso in cui le due particelle si trovano in stati diversi. Questo tipo di statistica fu utilizzato da Bose e da Einstein per spiegare la distribuzione di energia della radiazione ed è nota come statistica di Bose-Einstein. Le particelle che la seguono sono dette bosoni. Un altro modo ancora di distribuire le particelle si può ottenere inserendo un ulteriore postulato secondo cui due particelle identiche non possono mai occupare lo stesso stato quantico. Ovviamente questo postulato riduce le possibilità di suddividere le particelle tra gli stati a una sola. Il postulato non è altro che il principio di esclusione di Pauli, dal quale deriva la struttura a gusci di atomi e nuclei. La statistica basata su questo postulato è detta statistica di Fermi-Dirac e le particelle che ne seguono le regole sono chiamate fermioni.

Avviene (per motivi non ancora del tutto chiariti) che le particelle con spin

semintero siano fermioni, mentre le particelle con spin intero (0 o 1) siano bosoni. L'importante per noi è che la statistica, come lo spin, caratterizza la particella. Una particella di un certo tipo seguirà sempre la stessa statistica. Cosí talvolta si dice che vi è un principio di conservazione della statistica. Si può provare che la statistica di un sistema di particelle identiche è legata alla simmetria della funzione d'onda rappresentante il sistema. L'identità delle particelle introduce nel problema considerazioni di simmetria, poiché ci si può chiedere che cosa accada alla funzione d'onda quando s'interscambiano tutte le coordinate di due particelle identiche. Quando vi sono soltanto due particelle, la nuova funzione d'onda sarà identica a quella originaria, oppure muterà di segno. In quest'ultimo caso si dice che la funzione d'onda è antisimmetrica. Chiaramente una funzione d'onda di un sistema di bosoni non può essere antisimmetrica poiché la funzione non può eventualmente cambiare di segno quando si interscambino particelle con tutti i numeri quantici uguali. D'altra parte, la statistica di Fermi-Dirac e il principio di esclusione di Pauli si ricavano facilmente dall'assunto che la funzione d'onda di un sistema di elettroni (e altre particelle con spin 1/2) dev'essere antisimmetrica. Riepilogando, lo spin, la statistica e la simmetria della funzione d'onda, strettamente legati l'uno all'altro, dividono tutte le particelle in due gruppi esclusivi: bosoni e fermioni.

3.5. Creazione e annichilazione di particelle.

Sino ad ora si sono discusse due leggi di conservazione prive di controparte classica in relazione allo spin e alla statistica delle particelle. Si è visto anche come le leggi classiche di conservazione dell'energia e del momento debbano essere revisionate in conseguenza delle proprietà delle particelle della radiazione e delle regole di quantizzazione. Si vedrà ora che un'altra legge classica – la legge di conservazione della materia – dev'essere del tutto abbandonata in base alla meccanica quantistica. Si è scoperto che le particelle possono disintegrarsi completamente e che altre particelle possono crearsi al loro posto. Il primo ad essere scoperto è stato il cosiddetto processo di produzione di coppie, nel quale un fotone di raggi γ ad alta energia si annichila lasciando al suo posto un elettrone e un positrone. Fenomeni di questo tipo indeboliscono l'intero concetto di particella elementare. Come si è visto, dal tempo dei Greci le unità fondamentali della natura furono considerate indistruttibili, con la conseguenza che il loro numero in qualsiasi sistema chiuso doveva essere costante. Nella meccanica quantistica contemporanea non esistono tali unità indistruttibili; non vi sono particelle troppo «elementari» per essere create o annichilate. Ne segue che né il numero delle particelle in un sistema, né la massa del sistema sono costanti. Nel paragrafo dedicato alla teoria della relatività si sono già discussi la violazione della costanza della massa e il concetto unificato di massa-energia che obbedisce a una legge di conservazione. Ma una particella relativistica avrebbe ancora una sua identità: un elettrone sarà un elettrone di fronte a qualsiasi osservatore pur potendo essere piú pesante o piú leggero in funzione della velocità relativa dell'osservatore. Nella meccanica quantistica questa identità non viene mantenuta. La legge di conservazione dell'energia (massa-energia) è ritenuta valida e in ogni processo di creazione e annichilazione l'energia totale dei prodotti è uguale a quella delle particelle scomparse. Anche la legge di conservazione della carica è rigidamente conservata. Cosí nella produzione di coppie un positrone dev'essere creato con l'elettrone in modo che la loro carica totale sia zero come quella del fotone annichilato.

3.6. Particelle elementari.

Prima di continuare la discussione sulle leggi di conservazione è essenziale avanzare alcuni rilievi relativi alla classificazione delle particelle elementari. La fisica classica si occupa di due interazioni: l'interazione gravitazionale e l'interazione elettromagnetica. Nel dominio della fisica nucleare e della teoria delle particelle elementari furono scoperte due nuove interazioni: l'interazione forte e l'interazione debole. In confronto all'interazione forte, l'interazione elettromagnetica è 10^4 volte meno intensa, l'interazione debole è approssimativamente 10^{14} piú piccola e l'interazione gravitazionale è 10^{40} volte piú debole! Le particelle sono classificate in base al genere predominante della loro interazione tipica. È però possibile che una particella sia implicata contemporaneamente in diverse interazioni. Le particelle che possono interagire in modo forte sono dette adroni. Gli adroni si suddividono in due gruppi: bosoni-adroni, detti mesoni; e fermioni-adroni, detti barioni. Le particelle che interagiscono debolmente prendono il nome di leptoni e rientrano tutte tra i fermioni. L'interazione elettromagnetica si può descrivere come scambio del quanto del campo elettromagnetico (fotone) tra le cariche elettriche. In analogia con questa descrizione Yukawa propose di descrivere l'interazione forte come scambio di mesoni tra i nucleoni. Sono stati compiuti tentativi per attribuire una particella all'interazione debole e al campo gravitazionale, ma sinora non è stato possibile trovarne una conferma sperimentale.

3.7. Parità, inversione del tempo e coniugazione di carica.

Si è già notato che, ogniqualvolta un sistema è invariante per una data trasformazione, ne discende una legge di conservazione. Si dice che la trasformazione in quel caso è una *simmetria* del sistema. Una simmetria si può esprimere in termini «attivi», come trasformazione del sistema stesso, oppure in termini «passivi», come trasformazione del sistema di coordinate. La descrizione passiva è molto spesso piú opportuna, benché ovviamente il significato fisico sia identico nei due casi. Due altre trasformazioni nello spazio e nel tempo rivestono per noi importanza: la riflessione nello spazio e la riflessione nel tempo. La riflessione spaziale è descritta da un operatore cui si dà il nome di operatore di parità e la riflessione temporale da un operatore detto operatore di inversione del tempo.

3.8. Inversione dello spazio.

L'operatore di parità trasforma un vettore (x, y, z) nel vettore $(-x, -y, -z)$. Un sistema di coordinate destrorso ne verrà trasformato in un sistema di coordinate sinistrorso (un sistema di coordinate destrorso è un sistema che procede nella direzione di z positivo quando si passa dalla direzione di x positivo alla direzione di y positivo. Un sistema sinistrorso in questo caso avanza nella direzione di z negativo). Una variabile quantica può essere invariante o cambiare di segno rispetto all'inversione dello spazio. Ad esempio, l'operatore del momento lineare cambierà segno per l'inversione, mentre l'operatore del momento angolare resterà immutato. Si dice che uno stato quantico che inverte il segno per azione dell'operatore di parità ha una parità pari, mentre uno stato che rimanga inalterato è uno stato di parità dispari. Tutte le leggi della fisica classica sono invarianti per l'inversione dello spazio. Ciò significa che è impossibile operare una distinzione tra un processo e la sua riflessione nello specchio d'inversione dello spazio, in altre parole che è impossibile stabilire se nella descrizione del processo è stato usato un sistema di coordinate destrorso o sinistrorso. Si pensava che gli stessi principî valessero nel campo delle particelle elementari. Ad ogni particella si assegna una «parità intrinseca» (1 oppure -1) e la parità di un sistema si può calcolare se sono note le parità delle particelle che lo costituiscono. In generale l'attribuzione della parità intrinseca non è assoluta e ha significato soltanto in relazione ad altre particelle. Cosí un fermione e la sua antiparticella avranno parità opposte, mentre un bosone presenterà parità identica alla sua antiparticella. Ammesso che l'assegnazione sia fatta in modo coerente, la legge di conservazione della parità vale per le interazioni forti e per le interazioni elettromagnetiche. Ciò vuol dire che in ogni reazione forte o elettromagnetica la parità dello stato finale è uguale a quella dello stato iniziale. La scoperta che l'interazione debole non conserva la parità fu piú che sorprendente per la comunità dei fisici, poiché la conservazione della parità era quasi considerata un principio a priori. L'audace ipotesi che la parità potesse non conservarsi nelle interazioni deboli fu avanzata da Lee e Yang nel 1956 come spiegazione dell'esistenza di due interazioni di decadimento per il mesone K, interazioni che terminano con stati di parità differenti. Gli esperimenti proposti da Lee e Yang furono realizzati e dimostrarono che la violazione della conservazione della parità è invero tipica dell'interazione debole.

3.9. Inversione del tempo.

L'operatore di inversione del tempo può essere pensato come una riflessione dell'asse del tempo nell'origine, ossia $t \to -t$. Il significato di quest'operazione si può comprendere immaginando una pellicola cinematografica di un processo fisico fatta scorrere all'indietro. L'invarianza rispetto all'inversione del tempo esprime il fatto che il processo nella pellicola cinematografica sia probabile esattamente quanto il processo reale. Per illustrare l'assunto supponiamo di avere

un'interazione tra particelle elementari $A+B \to C+D$. Ogni particella si trova in uno stato di momento e spin definiti. Se l'inversione del tempo è una simmetria valida, allora la probabilità che quella reazione abbia luogo è identica a quella della reazione $C+D \to A+B$, in cui tutti i numeri quantici di momento e di spin hanno segno opposto. Nell'esperienza quotidiana ci si imbatte in molti processi che non sembrano invarianti rispetto all'inversione della direzione del tempo. Questi sono processi che coinvolgono un gran numero di particelle singole e sono regolati dalle leggi della meccanica statistica. Se però si prendono in esame processi individuali non implicanti una variazione di entropia, quali la rotazione di una particella, la collisione di due particelle, ecc., sembra naturale supporre che le leggi che governano tali processi siano assolutamente invarianti rispetto all'inversione del tempo. Tuttavia nel 1964 fu scoperta in un'interazione debole una prima violazione dell'invarianza rispetto all'inversione del tempo. I risultati erano soltanto leggermente deviati rispetto a quelli previsti nel caso di perfetta simmetria dell'inversione del tempo. Questi risultati sono tuttora aperti a due interpretazioni: o l'inversione del tempo è soltanto approssimativamente simmetrica nell'interazione debole, oppure è possibile un nuovo tipo d'interazione – una interazione superdebole – che infrange interamente il principio d'invarianza rispetto all'inversione del tempo.

3.10. Coniugazione di carica.

Un'altra simmetria che appare strettamente legata alle simmetrie di parità e d'inversione del tempo è la simmetria fondamentale tra particelle e antiparticelle. Si è stabilito sperimentalmente che per ogni particella esiste un'antiparticella con stessa massa e spin ma opposta carica elettrica (e altre cariche discusse in seguito). Una particella neutra viene talvolta considerata identica alla sua antiparticella, come si verifica col fotone, ma può altresí distinguersi da essa, come accade con il mesone K^0 e la sua antiparticella. Un modello ingegnoso sulla base del quale si possono prevedere le antiparticelle venne suggerito da Dirac nel 1931. Dirac costruí un'equazione d'onda relativistica che ammetteva soluzioni a energia negativa. Egli ipotizzò che tutti questi stati fossero occupati da elettroni in modo che non potessero verificarsi in essi delle transizioni, dal momento che un salto in uno stato occupato è impedito dal principio di esclusione di Pauli. Se un elettrone scompare dallo stato negativo, lascerà un «buco» e si può dimostrare che tale «buco» si comporterà esattamente come una particella avente la massa di un elettrone ma carica positiva. L'elettrone positivo profetizzato da Dirac venne effettivamente scoperto un anno dopo da Anderson e gli fu dato il nome di positrone. Come Dirac aveva previsto, si scoprí che la creazione del positrone si accompagna alla formazione di un elettrone e che il positrone ha vita assai breve poiché il «buco» tende a essere «riempito» nuovamente da un elettrone. Sino ad oggi tutte le antiparticelle previste sono state trovate (la teoria degli stati a energia negativa di Dirac non sempre viene però ritenuta valida). La simmetria tra particelle e antiparticelle dovrebbe trovare la propria espressione nel formalismo della meccanica quantistica. Ciò si ottiene definendo un operato-

re, denominato operatore di coniugazione di carica, che trasforma lo stato di una particella nello stato della sua antiparticella, ossia inverte la carica ma non lo spin e il momento. Invarianza rispetto alla coniugazione di carica significa che per ogni interazione tra particelle deve esistere un'equivalente interazione tra le antiparticelle. Le interazioni elettromagnetica e forte sono invarianti rispetto alla coniugazione di carica, mentre l'interazione debole non lo è. Questo fece sorgere una difficoltà per la previsione di antiparticelle delle particelle a interazione debole. Si è però avanzata l'ipotesi che benché la parità, l'inversione del tempo e la coniugazione di carica non siano simmetrie dell'interazione debole, la trasformazione combinata di tutte e tre è una simmetria di questo tipo e tutte le interazioni note sono invarianti per essa. La trasformazione combinata viene di solito definita TCP (T per l'inversione del tempo, C per la coniugazione di carica e P per la parità). Sinora non è stata individuata alcuna violazione dell'invarianza del TCP.

3.11. Isospin.

Le particelle a interazione forte, gli adroni, sembrano riunite in gruppi, i cui membri hanno stessi spin, parità intrinseca e numero barionico (cfr. oltre). I membri di ognuno di questi gruppi hanno cariche elettriche diverse e piú o meno la stessa massa. L'interazione forte sembra avere la stessa intensità tra membri diversi di un gruppo di questo genere. Ad esempio il protone e il neutrone hanno una differenza di massa di circa 0,1 per cento, uguali spin e parità e differiscono per la carica elettrica (il neutrone è neutro, mentre il protone ha carica positiva). L'interazione forte avrà la stessa intensità per tutte le coppie: protone-protone, protone-neutrone e neutrone-neutrone. Alla luce di queste considerazioni si suggerí di considerare il protone e il neutrone come due stati della medesima particella, detta nucleone, e di spiegare la lieve differenza nella massa dei due stati come risultante dall'energia delle interazioni elettromagnetica e debole, capaci di operare una distinzione. Il fatto che l'interazione forte non discrimini tra stati nucleonici o tra altri membri di gruppi analoghi può essere visto come una manifestazione di una nuova simmetria. Quando si mise a punto il formalismo di questa nuova simmetria ci si accorse che era strettamente analogo al formalismo del momento angolare e dello spin. In quel formalismo si definisce un nuovo operatore detto operatore dell'isospin (o spin isotopico), indicato con I. I ha tre componenti e, quando commuta con la hamiltoniana, I e una delle sue componenti, poniamo I_3, sono costanti di moto. Cosí come la conservazione dello spin e del momento angolare si può pensare risultante da una simmetria rotazionale di un sistema nello spazio reale, allo stesso modo la conservazione dell'isospin è conseguenza di una simmetria rotazionale di un sistema in uno spazio astratto denominato spazio dell'isospin o spazio delle cariche. Il cambiamento di direzione nell'isospazio ha il significato fisico di una variazione della carica elettrica del sistema e l'invarianza dell'interazione forte per una tale rotazione indica che essa è indipendente dalla carica elettrica. Ovviamente l'interazione elettromagnetica, essendo funzione delle cariche delle particelle interagenti, non

conserva l'isospin. Anche nell'interazione debole l'isospin non si conserva, ma la violazione è pur essa soggetta alla legge: si è scoperto sperimentalmente che la differenza nel valore di I_3 tra gli stati iniziale e finale di una reazione obbedisce alla regola $|\Delta I_3| = 1/2$. Nel caso dell'interazione forte, ove l'isospin si conserva, per caratterizzare uno stato quantico si possono usare i numeri quantici di isospin, analoghi ai numeri quantici di spin. In questo caso si applicano le limitazioni imposte dal principio di esclusione di Pauli, in altre parole la funzione d'onda totale dev'essere antisimmetrica (per i fermioni) o simmetrica (per i bosoni). Queste restrizioni spiegano perché siano stabili certi sistemi piuttosto che altri. Inoltre la legge di conservazione dell'isospin esclude il verificarsi di reazioni che non conservano l'isospin, e quindi fornisce regole selettive per le reazioni consentite.

3.12. Leggi di conservazione di quantità affini alla carica.

La legge di conservazione della carica elettrica, valida nella fisica moderna tanto quanto lo era nella fisica classica, differisce dalle leggi di conservazione dell'energia e del momento, e dalle invarianze di parità e inversione del tempo, in quanto non deriva da proprietà di simmetria dello spazio e del tempo. Essa non è una legge di conservazione *dinamica*. La legge di conservazione della carica elettrica richiede, come si è visto, che un'interazione in un sistema chiuso abbia termine con uno stato di carica uguale a quella dello stato iniziale. Poiché la carica elettrica è sempre un numero intero dell'unità fondamentale – la carica dell'elettrone – è possibile misurare la carica di un sistema di particelle elementari in valori interi 0, 1, 2... o -1, -2... Lo studio sperimentale delle particelle elementari ha dimostrato che vi sono altre proprietà *interne* (contrapposte a dinamiche) conservate di un sistema che sotto molti aspetti sono analoghe alla carica elettrica.

3.13. Numero barionico e numero leptonico.

Si è scoperto che in ogni interazione in cui intervengano dei barioni il numero totale dei barioni è costante. Se ad esempio una reazione inizia con due barioni, anche lo stato finale deve contenerne esattamente due. (Naturalmente può contenere altre particelle diverse dai barioni). Si dimostrò utile incorporare questa legge di conservazione nel formalismo della meccanica quantistica assegnando ad ogni barione un numero quantico additivo $B = 1$ e ad ogni antibarione un numero $B = -1$. Il numero barionico di un sistema corrisponde alla somma algebrica dei numeri barionici delle particelle componenti. Il numero barionico si conserva in tutte le reazioni sinora osservate. Il senso di questa legge è che essa impone limitazioni alle reazioni delle particelle elementari e mette in grado di prevedere quali siano possibili. In altre parole, tra tutte le reazioni permesse dalle altre leggi di conservazione, avranno luogo in pratica soltanto quelle che conservano il numero barionico.

Una legge di conservazione simile è quella del numero leptonico, definito

per i leptoni in analogia con il numero barionico. In realtà per i leptoni si conservano separatamente due quantità: il numero di elettroni e elettron-neutrini (con relative antiparticelle), denotato L_e, e il numero di muoni e muon-neutrini (con relative antiparticelle), denotato L_μ: entrambe le cariche leptoniche si conservano in tutte le interazioni. Un esempio potrebbe servire a illustrare il significato della legge: il muone negativo μ^- decade attraverso un'interazione debole in un elettrone e^-, un elettron-antineutrino $\bar{\nu}_e$ e un muon-neutrino ν_μ:

$$\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu.$$

La carica elettrica è -1 in entrambi i lati della reazione, L_μ vale 1 e L_e vale 0 da entrambe le parti. Un analogo processo $\mu^- \rightarrow e^- + e^- + e^+$ è proibito, benché si conservi la carica elettrica, in quanto le cariche leptoniche non si conservano.

3.14. Stranezza e ipercarica.

Ci si è resi conto che tutte le leggi di conservazione sinora trattate non sono sufficienti a spiegare i dati empirici relativi alle interazioni di particelle elementari. In altre parole vi sono reazioni ammesse da tutte le considerazioni precedenti, e che pure è impossibile rilevare. Queste reazioni sono soprattutto interazioni di decadimento di tipo debole di particelle prodotte attraverso interazioni forti. I processi di formazione di queste particelle sono anch'essi soggetti ad alcune limitazioni; ad esempio la particella Σ^+ si crea sempre con un mesone K^+ e mai con un mesone K^- – processo ammesso dalle leggi sinora prese in esame. Poiché il decadimento forte sembra impossibile per parecchie particelle, esse sono relativamente stabili, ovvero hanno una vita piú lunga del previsto. A causa del loro comportamento inspiegato quelle particelle furono chiamate *particelle strane*. Come si è visto, tutte le leggi di conservazione impongono restrizioni ai processi possibili; cosí si propose di spiegare l'impossibilità delle reazioni indicate presupponendo un'ulteriore legge di conservazione: la conservazione della stranezza. Il procedimento fu analogo a quello usato per altre leggi di conservazione di quantità affini alla carica. Un nuovo numero quantico, la stranezza S, fu assegnato alle particelle strane, che possono assumere valori interi positivi e negativi. Una particella e la sua antiparticella hanno stranezza opposta. La stranezza si conserva assolutamente nelle interazioni forti ed elettromagnetiche ed è violata nell'interazione debole in base alla regola $|\Delta S| = 1$. Per tutte le interazioni la stranezza mette quindi a disposizione un'altra regola di selezione, utilizzabile per predire le reazioni possibili. Spesso al posto della stranezza S si fa uso di un altro numero quantico detto ipercarica e contrassegnato con Y. La sua espressione è $Y = B + S$. Ne deriva che anche l'ipercarica si conserva nelle interazioni forti ed elettromagnetiche e che la sua violazione nell'interazione debole è soggetta alla regola $|\Delta Y| = 1$. L'ipercarica si lega alla carica elettrica Q e alla componente I_3 dell'isospin mediante l'equazione $Q = I_3 + (Y/2)$ e si dimostra piú conveniente della stranezza come numero quantico atto a classificare particelle.

3.15. Gruppi e simmetrie.

Si sono compiuti vari tentativi per classificare le particelle elementari in modo da chiarire la relazione tra i vari numeri quantici ad esse associati e spiegare le differenze di massa tra le particelle. Il formalismo matematico usato per tali classificazioni è la teoria dei gruppi, ma la sua presentazione va ben al di là dell'obiettivo di questo articolo. Ci si limiterà a fare alcune osservazioni.

La prima classificazione delle particelle in termini di gruppi di simmetria fu suggerita dal formalismo dell'isospin. Si è già visto che è possibile ordinare le particelle in multipletti di isospin diversificantisi soltanto per le cariche e che le leggere differenze di massa all'interno dei multipletti si possono approssimativamente spiegare come risultanti da interazioni che non conservano l'isospin. Le trasformazioni che generano le rotazioni nello spazio reale o nell'isospazio appartengono a un gruppo detto gruppo unitario speciale, di ordine 2, con sigla SU(2) e si possono rappresentare con matrici unitarie 2×2. L'invarianza per questo gruppo di trasformazioni si traduce nella conservazione dell'isospin. Si è tentato di indagare sulle rotazioni in uno spazio astratto tridimensionale, determinate dal gruppo unitario speciale di ordine 3, SU(3), e di disporre le particelle elementari in multipletti, detti di carica-ipercarica, piú ampi dei multipletti di isospin. Questo schema fu proposto da Gell-Mann e Ne'eman e consentí la previsione di una nuova particella – l'Ω^-. Contrariamente alle attese, SU(3) non è una simmetria perfetta neppure per l'interazione forte; cioè l'interazione forte non è invariante rispetto a rotazioni nello spazio dell'ipercarica e le differenze di massa all'interno dei multipletti di ipercarica non si possono spiegare con quello schema. Esistono altri gruppi di simmetria che sono stati proposti per la descrizione di proprietà invarianti di particelle elementari, ma nessuno di loro appare sino a questo momento soddisfacente e non si è scoperta alcun'altra legge di conservazione simile alla conservazione dell'isospin.

Per concludere si dovrebbero aggiungere alcune considerazioni: *a*) Sembra esservi una connessione tra l'intensità di un'interazione e il numero di simmetrie che possiede. Nell'interazione forte tutte le leggi di conservazione cui si è accennato sono valide. L'interazione elettromagnetica non conserva l'isospin, ma mantiene le altre simmetrie, e nell'interazione debole l'isospin e la stranezza sono violati e anche l'inversione dello spazio, l'inversione del tempo e la coniugazione di carica non costituiscono separatamente buone simmetrie. Si può quindi affermare che l'interazione è tanto piú simmetrica quanto piú è forte. *b*) Le leggi di conservazione svolgono nella fisica contemporanea un ruolo molto piú importante che non nella fisica classica. Il motivo è che nella fisica classica le equazioni del moto sono note e le leggi di conservazione non aumentano le informazioni fisiche, ma semplicemente rendono possibili nuove intuizioni sulla struttura della teoria e sulle proprietà della materia, dello spazio e del tempo. Nella meccanica quantistica la situazione è diversa. La hamiltoniana del sistema non è in generale nota e tutte le conoscenze fisiche sovente nascono dalle leggi di conservazione e da considerazioni di simmetria. L'ideale per cui la fisica do-

vrebbe essere basata su leggi di conservazione, che fa parte della nostra immagine della scienza, è da una parte il risultato dell'ignoranza di altri principî e dall'altra la conseguenza del notevole successo di un metodo estremamente elegante. *c*) La disamina delle leggi di conservazione classiche ha dimostrato che un concetto scientifico è in una condizione di «flusso» sinché viene incorporato in una legge matematica, e che dopo la sua scoperta è proprio questa legge a fornire la definizione del concetto. Nella fisica del XX secolo la situazione è un po' diversa. Poiché tutte le ricerche fisiche sono realizzate nell'ambito del formalismo matematico della meccanica quantistica, lo stadio di «concetto in divenire» è spesso breve o inesistente. Tuttavia la discussione che precede fornisce un gran numero di esempi di concetti privi di significato senza le leggi di conservazione loro inerenti. L'unico scopo dell'introduzione di concetti quali la stranezza o la carica dei leptoni consiste nell'esprimere una legge di conservazione, e senza tali leggi di conservazione i concetti sono del tutto trascurabili.

4. *Aspetti filosofici.*

4.1. Conservazione e causalità.

Come si è visto, il principio che la causa debba uguagliare l'effetto venne spesso invocato per appoggiare alcune delle leggi di conservazione classiche. Una breve discussione del legame tra leggi di conservazione e principio di causalità è quindi pertinente. Secondo la filosofia positivista della scienza il principio di causalità s'identifica col principio per cui tutti i fenomeni della natura sono soggetti a legge. Fornire una spiegazione causale per un evento A che si verifica per un insieme di condizioni B significa dimostrare che A è soltanto un caso di una legge generale per cui si verifica un evento di tipo A ogni volta che esistono condizioni di tipo B. Emile Meyerson, in *Identité et réalité* (1908) avanzò obiezioni su questa riduzione positivista della causalità. Dimostrare che l'accadere di A per le condizioni B è un caso di una legge generale, sostiene Meyerson, non basta a soddisfare la necessità di una spiegazione causale. Una spiegazione causale è quella che prova che, date le condizioni B, *deve* verificarsi l'evento A. In altre parole dare una spiegazione causale è fornire ciò che Leibniz definí una *ragione sufficiente*. Ora, Meyerson si chiede quale possa essere il motivo determinante per ciò che è condizionato nel tempo, e risponde che soltanto uno è possibile: quello della preesistenza. Le cose sono cosí perché erano già cosí in precedenza. Il principio di causalità esige quindi, secondo Meyerson, *identità nel tempo*. In questa forma il principio di causalità è piuttosto analogo al principio logico d'identità, salvo il fatto che il principio di causalità, a differenza dell'altro, non si riferisce a un'identità nel mondo statico astratto della logica, ma piuttosto a un'identità nel mondo dinamico reale, del quale il tempo è una dimensione. Ne discende che le uniche leggi scientifiche che soddisfino pienamente i requisiti del principio di causalità sono le leggi della conservazione. Questo è il motivo per cui le leggi di conservazione occupano il massimo rango nella gerarchia delle

leggi scientifiche: esse sono principî esplicativi nel senso piú profondo. Sebbene l'analisi di Meyerson dia una spiegazione storica dello status delle leggi di conservazione nell'immagine della scienza sul finire del XIX secolo, essa non può essere accolta come un'argomentazione filosofica o psicologica conclusiva a favore della priorità delle leggi di conservazione. Le leggi di conservazione forniscono soltanto un modello statico del mondo. Se le cose sono tali poiché furono cosí in passato non vi è alcuna ragione sufficiente per cui B preceda A e non viceversa. In altre parole le leggi di conservazione non provvedono una *direzione* per l'attualizzarsi dei fenomeni. Vi sono altre leggi, non di conservazione, ma capaci di fornire una spiegazione dell'elemento dinamico dei fenomeni. Un esempio di legge di questo tipo è la seconda legge della termodinamica. Vi sono anche leggi di conservazione che non paiono soddisfare la necessità di una spiegazione causale in quanto hanno troppo dell'ipotesi ad hoc, tagliata su misura per adattarsi ai dati osservati. Tale è ad esempio la legge di conservazione della stranezza che è stata discussa in precedenza.

4.2. Lo status epistemologico delle leggi di conservazione.

Nella fisica contemporanea le leggi di conservazione hanno lo stesso rango epistemologico di altre leggi scientifiche, ad esempio sono ipotesi da corroborare con l'esperienza. Non sempre è stato cosí. Alcune delle leggi di conservazione classiche spesso furono accettate come affermazioni a priori. Kant, ad esempio, dà una dimostrazione a priori della legge di conservazione della materia sulla base del suo principio metafisico secondo cui la sostanza non può essere né creata né distrutta. Analogamente, Herbert Spencer sostiene che siamo incapaci di concepire la creazione o l'annichilazione della materia, e che questa impossibilità è una cognizione aprioristica del massimo livello. Spencer avvertí la necessità di presentare una spiegazione per il fatto storico che un principio «a priori» fu stabilito soltanto nel XVIII secolo e che prima d'allora l'umanità sembrava alquanto desiderosa di pensare «l'impensabile». Ma la spiegazione da lui offerta non è – per non dire altro – molto convincente. Anche la legge di conservazione dell'energia, poco dopo la sua scoperta, fu indicata come principio a priori. Come si è visto persino Joule, il cui lavoro *sperimentale* tanto contribuí all'affermazione del principio, riteneva che conclusioni simili alle sue si potessero ricavare su base aprioristica. Nel corso della nostra generazione il mondo accademico fu sconcertato dalla confutazione della legge di conservazione della parità e dalla piú recente indicazione che i processi fisici non sono invarianti rispetto all'inversione del tempo. L'inversione dello spazio e l'inversione del tempo sembravano scaturire dagli assunti intuitivi sullo spazio e sul tempo in modo cosí naturale da essere quasi considerate come principî a priori. La confutazione di queste leggi, e di altre leggi di conservazione quali la legge di conservazione della materia, dimostra che le leggi di conservazione, come altre leggi scientifiche, devono essere sottoposte a prove sperimentali e che i tentativi di ricavarle da considerazioni aprioristiche sono futili. Di piú, la conclusione che le leggi di conservazione non siano leggi a priori poggia sullo stato epistemologico dei nostri concetti dello spazio

e del tempo. Come si è visto, le simmetrie di inversione dello spazio e inversione del tempo si fondano sulle nostre concezioni dello spazio e del tempo, concezioni che sembravano tanto intuitive da essere considerate valide a priori. Tuttavia, la negazione delle leggi di conservazione basate su queste simmetrie rinforza l'affermazione (sostenuta anche da altri argomenti) che i nostri concetti di spazio e tempo abbiano anche natura a posteriori.

D'altra parte si è spesso argomentato che le leggi di conservazione sono semplicemente generalizzazioni induttive basate sull'osservazione che certi processi non sembrano mai avvenire. Ad esempio è stato detto che la legge di conservazione dell'energia è una conclusione induttiva ricavata dall'impossibilità di costruire una macchina realizzante il moto perpetuo. Si è già visto come questa affermazione sia lontana dalla verità storica, e come la convinzione che «qualcosa» si conservi in relazione alla materia, al moto o alla produzione di lavoro, abbia avuto un ruolo importante nella scoperta delle leggi di conservazione classiche. Tale persuasione, non ben definita ma pur tuttavia molto forte, superava di gran lunga i dati osservati. Si deve perciò concludere che le leggi di conservazione sono a posteriori, nel senso che debbono essere sperimentalmente valutate e non possono essere derivate per via logica, ma che la loro scoperta è il risultato di una fertile interazione tra esperienza e metafisica.

4.3. Le leggi di conservazione e la psicologia evolutiva di Jean Piaget.

Jean Piaget e la sua scuola gettarono le basi per lo studio psicologico dello sviluppo concettuale nei bambini. Piaget per primo richiamò l'attenzione sull'importanza della formazione dei concetti di conservazione per lo sviluppo mentale del fanciullo. Il lavoro di Piaget prende le mosse dalla tesi di Meyerson secondo cui le leggi di conservazione sono estremamente importanti per la comprensione intellettuale del mondo. Poiché è impossibile dare in questa sede una esposizione soddisfacente dell'opera di Piaget, si limiterà la discussione ad alcuni punti fondamentali. Lo sviluppo dei concetti di conservazione, secondo Piaget, è uno stadio indispensabile nell'evoluzione intellettuale del bambino. La mancanza di questi concetti o una loro comprensione tarda e incompleta può indicare un certo grado di ritardo mentale. Le leggi di conservazione cui fa riferimento Piaget non sono naturalmente quelle discusse in precedenza. È ovvio che un bambino che non scopra la legge di conservazione della carica elettrica non sia mentalmente ritardato. Piaget ha studiato la conservazione della quantità di materia, del peso, del volume e del numero; egli allude anche a un concetto da lui definito «conservazione dell'oggetto», indicante la realizzazione da parte del fanciullo che gli oggetti non cessano di esistere quando egli smette di osservarli, ossia la realizzazione che gli oggetti hanno un'esistenza indipendente. Le trasformazioni per cui il bambino scopre che le quantità citate (peso, volume, ecc.) si conservano sono di tipo molto semplice, come la variazione di forma di una pallina di creta, o il versamento di acqua in contenitori di forma diversa. Lo scopo delle leggi di conservazione scoperte dal fanciullo è molto limitato. Vi sono molte trasformazioni che non modificano il volume di un oggetto, ed esistono

trasformazioni che cambiano la massa e il peso. Non vi è quindi analogia tra il *contenuto* dei concetti di conservazione del bambino e le leggi di conservazione scientifica. Tuttavia vi è una somiglianza tra la struttura teorica che sviluppa il bambino per inquadrare le sue esperienze e la struttura teorica della scienza: in entrambe le leggi di conservazione sembrano infatti svolgere un ruolo importante.

Secondo Piaget, l'acquisizione di tutte le leggi di conservazione da parte del bambino segue lo stesso schema: il bambino passa dal primo stadio, in cui non ha concetti di conservazione, attraverso un secondo stadio, nel quale si manifesta una conflittualità tra conservazione e non-conservazione, al terzo stadio, quello finale, in cui raggiunge una legge di conservazione quasi assiomatica e tale da non potersi piú estinguere anche quando il bambino si trova di fronte a una prova contraria. Piaget fa vedere che i *concetti* di peso, volume, ecc. sono afferrati dal bambino soltanto quand'egli perviene allo stadio finale e acquisisce la legge di conservazione relativa. Piú specificamente: nel primo e nel secondo stadio un bambino potrebbe ancora pensare che una pallottola di argilla sia piú pesante della «salsiccia» ottenuta con lo stesso materiale. Ciò non perché, sostiene Piaget, il bambino sappia cos'è il peso e non conosca la sua proprietà di conservarsi, ma piuttosto perché non ha ancora acquisito il *concetto* di peso. La realizzazione che il concetto di peso deve essere separato da quello di forma o lunghezza o «grossezza» viene soltanto dalla consapevolezza che il peso, a differenza della forma o della lunghezza o della «grossezza», si conserva. Vediamo cosí che la teoria dei «concetti in divenire» è valida tanto per il singolo bambino quanto per il processo di formazione dei concetti nella scienza.

Un altro punto interessante è il legame tra leggi di conservazione e atomismo. Piaget ha scoperto che se il bambino ha una concezione «atomistica» della materia potrà piú facilmente acquisire una legge di conservazione del peso e della quantità di materia. Per quel bambino la quantità di materia è il numero degli «atomi» ed è probabile che egli capisca che una variazione della forma di un oggetto lascia invariante il numero di «atomi». Si è già osservato che anche in campo scientifico l'atomismo fu la base per la prima versione della legge di conservazione della materia.

Nonostante il fatto che sembri sussistere qualche somiglianza tra lo sviluppo intellettuale del bambino e l'evoluzione della scienza, non se ne dovrebbero trarre conclusioni affrettate. Il fatto che il bambino normale impari varie leggi di conservazione di portata limitata non implica che la scienza *debba* essere fondata sulle leggi di conservazione. La storia della scienza dimostra che la liberazione dalle concezioni del mondo frutto del buon senso è una condizione necessaria per il progresso della scienza. Non è possibile concludere dunque che ciò che fa parte del senso comune debba essere indispensabile nella scienza.

Inoltre bisogna ricordare che gli esperimenti di Piaget sono dipendenti dalla cultura. È incerto se un bambino svilupperebbe «spontaneamente» concetti di conservazione, se non fosse inserito in una società che li dà per scontati.

4.4. Mutamento e persistenza riesaminati.

A conclusione dell'articolo è il caso di riprendere in esame il problema del mutamento nei confronti della persistenza, interrotto nella prima parte, in considerazione degli ultimi sviluppi nella scienza. Qual è l'atteggiamento dei fisici moderni verso la dicotomia parmenideo-eraclitea?

Il cambiamento, secondo la fisica contemporanea, non è arbitrario. Non tutti i processi sono possibili. Come si è visto, vi sono certe regole di selezione basate su leggi di conservazione che prevedono quali processi siano possibili e quali no. Inoltre, la fisica considera certe *strutture* come costanti. Un dato sistema (ad esempio un atomo d'idrogeno) avrà sempre gli stessi possibili stati quantici.

D'altra parte, non è piú vero che «nulla si crea e nulla si distrugge». Le particelle elementari si creano e annichiliscono continuamente. Inoltre, un sistema dove un certo numero di elettroni «riempia» un certo numero di stati quantici è *indistinguibile* da un sistema nel quale gli elettroni abbiano «cambiato posizioni». Le particelle, secondo il nostro modello del mondo, mancano d'individualità e di permanenza nel tempo. Si potrebbe forse parafrasare un famoso detto di Eraclito e dire che secondo la fisica quantistica è impossibile colpire due volte lo stesso elettrone. L'identità non ha significato in relazione alle particelle elementari. Tuttavia è possibile riempire due volte lo *stesso* stato quantico. Gli stati hanno un'individualità di cui sono prive le particelle. Si potrebbe perciò affermare che la fisica attuale è parmenidea per quanto riguarda la struttura, ma eraclitea per quanto concerne la materia. [Y. E. e Y. B. M.].

Strettamente legate al problema della comprensione del mutamento nella **scienza** e nella **filosofia**, le leggi fisiche di conservazione, parificate epistemologicamente alle altre leggi scientifiche (cfr. **legge**) in quanto soggette alla verifica sperimentale (cfr. **verificabilità/falsificabilità**, **esperimento**), stabiliscono quali proprietà di un **sistema** isolato siano invarianti nel tempo. La **fisica** classica individua i principî di conservazione della **materia**, a partire dalla definizione di massa, di momento angolare e lineare (cfr. **moto**), di carica elettrica, di **energia** (cfr. **forza**), con le relative implicazioni circa l'impossibilità di costruire una **macchina** in moto perpetuo. Facendo uso di un formalismo matematico assai complesso (cfr. **strutture matematiche**), la fisica contemporanea ha poi scoperto molte nuove leggi di conservazione, nel quadro della teoria della **relatività**, della teoria dei **quanti** e della meccanica quantistica, che attribuisce alle varie particelle atomiche (cfr. **atomo e molecola**, **particella**) delle **funzioni** d'onda di tipo probabilistico (cfr. **probabilità**, **distribuzione statistica**). Il concetto di conservazione/invarianza pone infine problemi di ordine piú generale, in relazione sia al legame con il principio di causalità (cfr. **causa/effetto**), sia al ruolo svolto nella formazione delle funzioni cognitive (cfr. **cognizione**, **concetto**) nello sviluppo intellettuale del bambino (cfr. **cervello**, **mente**).

Entropia

Il concetto di entropia è stato introdotto nella fisica per rendere ragione di un gran numero di fenomeni che sembrano manifestare una qualche direzionalità e che la meccanica classica non riusciva a spiegare. Si osserva che il calore fluisce sempre da un corpo caldo verso l'ambiente piú freddo, e mai in senso contrario. Una bottiglia di profumo in un angolo di una stanza spande il suo odore dappertutto, ma non si è mai assistito al graduale riempimento di una bottiglia vuota posta in una stanza invasa da profumo. Si può distruggere una parete di mattoni facendo esplodere nei pressi un po' di dinamite, ma nessuno ha mai proposto di costruire una parete per deflagrazione di un mucchio di mattoni. Questi fatti sono talmente familiari da apparire ovvi, ma, se si tenta di spiegarli in base alla meccanica classica, ci si rende conto dell'impossibilità di dedurre l'orientamento dei fenomeni dalle equazioni del moto. Le equazioni della meccanica classica sono invarianti rispetto all'inversione del tempo e inutili ai nostri fini. La scienza della meccanica, capace di render conto con gran precisione dei movimenti dei pianeti, non riesce a spiegare alcuni dei fatti basilari dell'esperienza quotidiana. È necessario introdurre un nuovo principio fisico: il principio dell'incremento di entropia, o seconda legge della termodinamica. In base a questo principio è possibile spiegare la direzione dei fenomeni citati in termini di aumento di entropia.

Ma cos'è l'entropia? Difficile è darne una definizione operativa semplice. Come per molti altri concetti scientifici, la definizione di entropia dipende dalla teoria e ha subito significative modifiche con lo sviluppo della termodinamica, della meccanica statistica e della teoria dell'informazione [cfr. Fast 1959]. Il carattere comune di tutte le definizioni possibili è che l'entropia di un sistema fornisce una misura della capacità del sistema di subire trasformazioni spontanee. In questo articolo si presenteranno e discuteranno varie definizioni piú particolareggiate.

L'importanza del principio dell'aumento dell'entropia risiede nella sua estesa applicabilità. Si tratta di una legge che interessa tanto l'ingegnere quanto il fisico teorico, e i sistemi che le obbediscono spaziano da quelli fisici a quelli chimici e biologici. Inoltre, come si vedrà, si può introdurre il concetto di entropia in senso puramente astratto ogni qualvolta è in discussione una distribuzione di probabilità. È possibile applicare tale concetto astratto di entropia a una gamma ancora piú estesa di problemi, al di là dell'obiettivo della fisica, quali la linguistica, l'analisi letteraria, la codificazione dei linguaggi, la teoria della musica e molti altri.

Da un punto di vista matematico formale si osserva un'analogia sorprendente tra i diversi concetti di entropia. In senso concettuale i vari concetti si differenziano però in misura tale che è stata messa in dubbio la possibilità stessa di unificarli sotto lo stesso nome. Il concetto di entropia non è perciò soltanto un concetto fondamentale della scienza contemporanea, ma fornisce

anche un'esemplificazione assai illuminante dell'evolversi dei concetti scientifici. Lo sviluppo della scienza è, da un lato, un processo continuo, nel quale ogni stadio emerge dallo stato dei problemi degli stadi precedenti; dall'altro è un processo discontinuo, nel quale spesso le analogie formali mascherano cambiamenti radicali.

I primi tre paragrafi dell'articolo si occupano del concetto di entropia rispettivamente nella termodinamica classica, nella meccanica statistica e nella teoria dell'informazione. Nel quarto saranno dibattuti alcuni problemi filosofici collegati con tali argomenti.

1. *L'entropia nella termodinamica classica.*

Il termine 'entropia' fu coniato da Rudolf Clausius nel 1865, il quale lo ricavò dalla parola greca τροπή 'trasformazione'. Clausius intendeva quantificare le trasformazioni che si verificano nel carattere dell'energia di un sistema sottoposto a cambiamento termodinamico. Una presentazione dogmatica del concetto di entropia lo renderebbe scarsamente comprensibile; di conseguenza è preferibile seguire la linea storica e introdurlo soltanto dopo avere messo in evidenza il contenuto della seconda legge della termodinamica.

1.1. La seconda legge della termodinamica: Sadi Carnot.

I fondamenti della seconda legge della termodinamica furono gettati da un fisico e ingegnere francese, Sadi Carnot, nei suoi studi relativi al funzionamento dei motori termici. Un motore termico è un dispositivo in grado di produrre lavoro usando calore. Un esempio è costituito dai motori a vapore, che hanno avuto larga diffusione durante la rivoluzione industriale e sono stati costantemente migliorati e adattati a nuovi compiti. Tuttavia in pratica non esisteva una teoria scientifica delle loro modalità di funzionamento. Sadi Carnot, cercando di consolidare le carenti basi teoriche, si pose l'interrogativo seguente: è possibile aumentare indefinitamente il rendimento di un motore termico, o esiste un limite al rendimento, insuperabile in linea di principio? Può esistere un agente piú efficiente del vapore, usabile in un motore termico? In altre parole, il rendimento dipende dalla natura della sostanza attiva? È essenziale che un motore termico abbia un'uscita e un ingresso di calore? Come è influenzato il rendimento dalla temperatura della sorgente dalla quale si estrae il calore?

Per rispondere a tali quesiti Carnot propose il principio che oggi porta il suo nome e che rappresenta la prima versione della seconda legge della termodinamica: la produzione di lavoro da parte di un motore termico deve sempre essere associata a un flusso di calore da un corpo a temperatura piú alta a un altro corpo a temperatura piú bassa. Questo processo di passaggio di calore da un corpo caldo a uno piú freddo è descritto da Carnot come ristabilimento dell'«equilibrio del calore», o, servendosi di un termine piú comune, ripristino dell'equilibrio termico. Carnot riteneva anche valido l'opposto: è possibile ottenere lavoro ogni qualvolta esista una differenza di temperatura e si possa ripristinare l'equilibrio termico. L'esistenza di *due* sorgenti di calore caratterizzate da una diversità di temperatura è quindi una condizione sufficiente e necessaria per la produzione di lavoro mediante calore. Va osservato immediatamente che, in base a questo principio, è impossibile far muovere un veicolo estraendo calore dall'ambiente col quale esso si trova in equilibrio termico. Altrimenti non avremmo bisogno di combustibile per azionare navi o aeroplani o automobili: essi potrebbero ricavare tutto il calore necessario dal mare o dall'aria e spostarsi per un periodo indefinito. La ragione dell'inesistenza di siffatte navi o automobili non risiede nell'incapacità dell'ambiente di fornire una quantità di calore sufficiente al loro funzionamento. Ciò che manca è una differenza di temperatura, un flusso di calore da un corpo caldo a uno freddo.

Forte di questo principio, Carnot si accinse quindi a esaminare le condizioni di massimo rendimento di un motore termico. Poiché il passaggio di calore da una temperatura piú alta ad una piú bassa può sempre produrre lavoro, ragionò Carnot, un trasferimento di calore senza produzione di lavoro, quale si verifica quando un corpo caldo e uno freddo sono posti a contatto diretto, equivale a uno spreco del lavoro che la cessione di calore avrebbe potuto generare. Un motore termico ideale deve quindi tendere a evitare una trasmissione diretta di calore lungo una differenza di temperature, onde produrre tutto il lavoro in teoria ottenibile. In altre parole, nel motore termico ideale tutti i passaggi di calore si debbono realizzare tra temperature uguali e tutti i riscaldamenti o raffreddamenti nell'ambito del motore si debbono effettuare tramite cambiamenti di pressione-volume, senza flusso termico diretto. Questo motore ideale, secondo Carnot, non sarebbe affatto in grado di funzionare, in quanto il calore non fluirebbe nel motore senza una differenza di temperatura, ma la minima deviazione dall'ideale, la piú lieve diversità di temperatura sarebbero sufficienti ad azionare la macchina e al contempo a mantenere il rendimento vicino al valore massimo.

Fondandosi su queste considerazioni, Carnot descrisse il motore termico ideale. Il motore opera in cicli successivi, in modo che al termine di ciascun ciclo la macchina ritorni allo stato iniziale. Ogni ciclo è formato dagli stadi seguenti: 1) il motore riceve calore da un corpo A, di temperatura elevata T_1 (sfruttando la minima possibile differenza di temperatura); 2) il motore produce lavoro senza scambio termico, mentre la sostanza attiva si raffredda alla temperatura piú bassa T_2; 3) si trasferisce calore dal motore a un serbatoio freddo B (la cui temperatura è quanto piú vicina possibile a T_2); 4) si spende lavoro per riportare la sostanza attiva alla temperatura piú alta T_1, senza passaggio di calore. Il lavoro netto prodotto dal motore è la differenza tra quello guadagnato nella fase 2 e quello speso nella fase 4.

Un'osservazione cruciale per la comprensione del funzionamento dei motori termici, e fondamentale per la scienza della termodinamica in generale, è il collegamento, stabilito da Carnot, tra *reversibilità* del motore e rendimento.

Il ciclo ideale di Carnot è completamente reversibile. Tutti e quattro gli stadi si possono realizzare in ordine inverso, naturalmente a patto che si spenda lavoro anziché ricavarne. La reversibilità del motore è una conseguenza delle precauzioni prese per evitare un riscaldamento o raffreddamento determinato da scambi termici. Se si fosse avuto un flusso termico diretto lungo un gradiente di temperatura, il ciclo non sarebbe stato reversibile, poiché l'operazione non avrebbe potuto verificarsi in senso opposto senza apportare determinate modifiche al funzionamento del motore. Ne deriva che il motore piú efficiente è quello reversibile e che qualsiasi ciclo irreversibile è destinato a produrre meno lavoro di quello reversibile. Il miglior uso che sia possibile fare del calore si manifesta quindi in un processo reversibile. Naturalmente in pratica il processo reversibile ideale è inottenibile: non è possibile evitare completamente il contatto diretto tra corpi caldi e corpi piú freddi. L'analisi teorica suggerisce però rigorosamente come si possa approssimare la situazione ideale.

L'analisi di Carnot si fonda soltanto sugli aspetti piú generali e essenziali del motore termico. Nessun dettaglio di progettazione o materiale può quindi influenzarla. Dopo aver descritto il suo motore reversibile, egli fu in grado di dimostrare che nessun'altra macchina può avere un rendimento maggiore (ad esempio una macchina con una diversa sostanza attiva) e che nessun'altra macchina può compiere lo stesso lavoro con una minor quantità di calore. Egli dimostrò anche che il rendimento del motore è tanto piú elevato quanto maggiore è la differenza di temperatura tra le sorgenti calda e fredda. La prova della prima affermazione è semplice e costituisce un esempio per molte altre prove avanzate dai seguaci di Carnot: «Se esistesse un metodo di impiego del calore preferibile a quello di cui ci siamo serviti, in altre parole se fosse possibile ottenere dal calore, con qualunque processo alternativo, una maggiore quantità di energia motrice... si potrebbe allora, dirottando una parte di questa energia, effettuare un riciclo di calore dal corpo B al corpo A, ovvero dal refrigerante alla sorgente, e cosí ripristinare la situazione originaria in modo da dare il via a un'operazione esattamente simile alla prima, ripetendo il processo piú volte. Ne deriverebbe non soltanto il moto perpetuo, ma un'illimitata creazione di energia motrice senza consumo di calore o di altro agente di sorta. Una simile creazione è completamente contraria alle idee oggi accettate, alle leggi della meccanica e della fisica razionale; è inammissibile» [Carnot 1824, p. 11].

L'argomento citato è una *reductio ad absurdum*, ove l'assurdo è l'esistenza di un perpetuum mobile. Tuttavia l'attento lettore moderno non concorderebbe sul fatto che il processo descritto porti di fatto a un perpetuum mobile, ma si limiterebbe ad ammettere che fornisce un dispositivo capace di rimuovere calore da un serbatoio freddo senza spendere lavoro. Un dispositivo di questo genere non è un perpetuum mobile nel senso stretto del termine. Ad ogni modo, sappiamo che è impossibile costruire una macchina siffatta e l'argomentazione di Carnot è valida anche se il termine è usato leggermente a sproposito.

Com'è possibile spiegare questo abuso da parte di Carnot? Per rispondere bisogna domandarsi quale modello, quale quadro d'insieme avesse in mente Carnot nell'immaginare i suoi cicli ideali. La risposta la fornisce esplicitamente lui stesso: «In base a quanto stabilito possiamo correttamente confrontare l'energia motrice del calore con quella di una cascata – entrambe hanno un massimo invalicabile... che dipende dall'altezza e dalla quantità dell'acqua. Analogamente col calore... la forza motrice dipende dalla differenza di temperatura e dalla quantità di calore (calorico)» [*ibid.*, p. 15]. La teoria del calore riflessa in questa analogia è quella del calore considerato come sostanza. Secondo questa teoria, nel motore termico non si consuma in realtà calore, proprio come in una macchina che sfrutti una cascata non si consuma acqua: è la *cascata* che produce il lavoro. Le opinioni degli storici della scienza divergono sul concetto carnotiano di calore. È risaputo che Carnot nutriva seri dubbi sulla natura del calore e che col procedere degli anni si convertí alla teoria cinetica del calore (come rivelano i suoi scritti postumi). Alcuni storici credono di scorgere quest'ultima teoria nelle *Réflexions*, ma questa interpretazione, alla luce dell'analogia citata, sembra impossibile. Se la corrispondenza tra flusso di calore e caduta d'acqua fosse stata giustificata, il processo ipotetico di Carnot darebbe di fatto un perpetuum mobile. Per noi, che non accettiamo piú l'analogia, il risultato non è un perpetuum mobile, ma piuttosto un tipo diverso di macchina impossibile: una macchina che trasferisce calore da un corpo freddo ad uno caldo senza spendere lavoro. In pratica la costruzione di una simile macchina non è piú probabile di quella di un perpetuum mobile «reale». È quindi del tutto naturale che a questo tipo di macchina si sia dato il nome di «perpetuum mobile del secondo ordine».

L'analogia calore-acqua suggerita da Carnot non era casuale. All'epoca di Carnot era d'uso tra gl'ingegneri progettare motori ad acqua capaci di funzionare anche in senso *contrario*, ossia come pompe idrauliche. Il rendimento di una tale macchina era definito come rapporto tra la quantità d'acqua restituibile all'altezza originaria mediante il ciclo invertito e la quantità totale di acqua motrice. Si sapeva che il motore idraulico ideale era quello completamente reversibile. Una macchina con rendimento superiore a quella reversibile è un perpetuum mobile. Poiché la teoria piú diffusa sul calore era a quell'epoca la teoria calorica, è perfettamente naturale che Carnot abbia fondato la sua analisi sull'analogia calore-acqua. È abbastanza interessante notare che nella derivazione dei risultati nulla dipende realmente dall'analogia. Carnot (forse a causa dei suoi dubbi) espose l'argomento con tale prudenza che tutti i risultati sono fondamentalmente corretti, anche se ben presto il presupposto, o meglio il modello retrostante al presupposto, si dimostrò inadeguato.

1.2. Il conflitto tra la prima e la seconda legge della termodinamica.

Al suo apparire l'analisi di Carnot suscitò scarsa eccitazione. In realtà sarebbe quasi potuta finire in un cassetto, non fosse stato per l'ingegner Emile Clapeyron, che nel 1834 riformulò la teoria di Carnot in linguaggio matematico

(tra l'altro, Clapeyron si appoggiò piú vistosamente di Carnot alla teoria calorica del calore, fatto che impedí una corretta comprensione della teoria di Carnot).

Piú di vent'anni dopo la pubblicazione delle memorie di Carnot, nel 1847, James Joule, in una conferenza tenuta a Oxford, descrisse i risultati del suo lavoro sperimentale sulla trasformazione di lavoro meccanico ed elettrico in calore. Secondo Joule, è possibile *convertire* energia meccanica in calore e calore in energia meccanica. Calore e lavoro meccanico sono in certa misura equivalenti e si possono trasformare l'uno nell'altro. Una data quantità di lavoro meccanico produrrà sempre la stessa quantità di calore, e viceversa. Joule calcolò anche con grande precisione l'equivalente meccanico del calore: ogni caloria equivale a $4{,}187 \times 10^7$ erg di energia meccanica.

A quell'epoca Joule non sapeva nulla del lavoro di Carnot, ma Lord Kelvin, presente alla conferenza, fu profondamente impressionato dai risultati di Joule e immediatamente imbarazzato da quella che gli sembrava una contraddizione insuperabile: se Carnot era nel giusto, è impossibile trasformare completamente una quantità di calore in lavoro, poiché il rendimento del motore termico ideale è di gran lunga inferiore a 1. Inoltre l'analisi di Carnot riguarda la produzione di lavoro mediante calore come conseguenza del *flusso* di calore, mentre Joule interpretava la produzione di lavoro come trasformazione o *conversione* di calore in lavoro. Entrambe le teorie erano convalidate da dati sperimentali. Come si poteva riconciliarle?

Fu Rudolf Clausius a trovare per primo (nel 1850) il bandolo della matassa. Egli distinse tra i risultati di Carnot e il modello su cui si fondavano, accettando in pieno i primi ma rifiutando il secondo. In questo modo la contraddizione scomparve. Clausius riteneva che Joule avesse ragione a invocare la teoria cinetica del calore. Il calore non è una sostanza indistruttibile che obbedisce a una legge di conservazione separata, ma è una forma di energia, una manifestazione dell'energia cinetica delle particelle che costituiscono la materia. In quanto forma di energia, il calore può essere sia generato da, che trasformato in, altre forme di energia. Queste trasformazioni seguono la legge di conservazione dell'energia o *prima* legge della termodinamica. La produzione di lavoro mediante calore non esige quindi soltanto una variazione nella distribuzione del calore (come implicava l'analogia calore-acqua di Carnot), ma un vero *consumo* di calore, come comprese Joule.

D'altra parte Carnot non sbagliava nel far notare che il processo di conversione di calore in lavoro non può avere luogo se non sussiste una differenza di temperatura, e se una parte del calore non passa da una temperatura piú alta a una piú bassa. Non è possibile realizzare una completa conversione di una certa quantità di calore in lavoro, senza il processo di raffreddamento di una parte del calore. Joule aveva ignorato questo lato del problema. L'analisi di Clausius sulla produzione di lavoro mediante calore [1864-67] poggia quindi su entrambi i principî, di Joule e di Carnot: 1) in tutti i casi in cui si ottiene lavoro a partire da calore, si consuma una quantità di calore proporzionale al lavoro compiuto, e, viceversa, spendendo una determinata quantità di lavoro si produce un'uguale quantità di calore (principio di Joule, una delle versioni della prima legge della termodinamica); 2) in tutti i casi in cui si produce lavoro mediante calore, ha luogo un passaggio di calore da un corpo piú caldo a uno piú freddo (principio di Carnot, una delle versioni della seconda legge della termodinamica). Riepilogando, il consumo e la trasmissione di calore sono entrambi condizioni necessarie per la produzione di lavoro mediante calore.

1.3. La seconda legge della termodinamica in termini di entropia.

Negli anni seguenti Clausius continuò a sviluppare la sua teoria. Considerando tutte le trasformazioni calore-lavoro come trasformazioni di energia, egli tentò di scoprire una funzione che soddisfacesse ai seguenti requisiti: 1) la funzione dovrebbe porre in relazione i vari tipi di possibili trasformazioni energetiche in un sistema, ossia le trasformazioni di calore da una temperatura piú alta a una piú bassa e le trasformazioni di energia meccanica in calore ad una determinata temperatura; 2) la funzione dovrebbe essere una funzione termodinamica propria, ovvero dovrebbe caratterizzare lo stato termodinamico del sistema indipendentemente dalla via seguita dal sistema per raggiungere quello stato. Si può facilmente dimostrare che il calore non è di per sé una funzione termodinamica propria in questo senso: è possibile pervenire allo stesso stato termodinamico (ad esempio uno stato caratterizzato dalla medesima temperatura, pressione e volume) sia aggiungendo calore al sistema sia compiendo su di esso un lavoro; il contenuto termico varierà di conseguenza in relazione al metodo usato e non costituirà una funzione univocamente definita dello stato termodinamico.

Clausius definí *entropia* la funzione che si dimostrò capace di soddisfare questi requisiti. L'entropia S di un sistema è espressa in modo differenziale. Per un processo *reversibile* infinitesimale vale la relazione:

$$dS = \frac{dQ}{T}.$$

La variazione di entropia in un processo reversibile infinitesimale è cioè data dal rapporto tra la variazione del contenuto termico del sistema, dQ, e la temperatura assoluta T propria del processo. (Per convenzione, dQ è positivo quando il sistema assorbe calore e negativo quando il sistema perde calore cedendolo all'ambiente). Per ogni cambiamento finito reversibile dallo stato A allo stato B, la variazione di entropia, ΔS, è data dall'integrale:

$$\Delta S = S_B - S_A = \int_A^B \frac{dQ}{T}.$$

Si noti che qui non è definito il valore assoluto dell'entropia, ma soltanto la sua *variazione*. La terza legge della termodinamica, o legge di Nernst, che attribuisce alle sostanze (in realtà soltanto alle sostanze idealmente cristallizzate) un'entropia nulla in corrispondenza dello zero assoluto, consente il cal-

colo dell'entropia assoluta. Si osservi anche che le variazioni di entropia sono definite *soltanto per processi reversibili*. Poiché l'entropia di un sistema è una funzione termodinamica propria, essa dipende soltanto dallo stato termodinamico del sistema. È quindi possibile trovare la variazione di entropia in un processo *qualunque*, reversibile o irreversibile, valutando la variazione di entropia che si sarebbe verificata in un processo reversibile avente gli stessi stati iniziale e finale.

Data questa definizione, è possibile riformulare la seconda legge della termodinamica in linguaggio matematico e in termini di entropia. Clausius ha dimostrato che si può esprimere la seconda legge con la seguente disuguaglianza (il segno $\oint$ indica un integrale lungo un ciclo):

$$\oint \frac{dQ}{T} \leq 0.$$

In ogni ciclo reversibile vale la relazione $\oint(dQ/T)=0$, mentre in ogni ciclo irreversibile (tutti i cicli reali sono di fatto irreversibili) vale $\oint(dQ/T)<0$. Sono impossibili cicli per cui $\oint(dQ/T)>0$. Tale disuguaglianza, talvolta definita disuguaglianza di Clausius, manifesta lo stesso principio contenuto nella formulazione precedente: l'impossibilità di un passaggio spontaneo di calore da un corpo freddo a uno caldo, o, in altre parole, l'impossibilità di una completa trasformazione in lavoro di una determinata quantità di calore. Per dimostrarlo, si supponga di riuscire a trasferire una quantità di calore Q' da un serbatoio a temperatura T a un serbatoio a temperatura *piú elevata*, T'. In base alla convenzione di segno, ciò significherebbe che:

$$\oint \frac{dQ}{T} = \frac{Q'}{T} - \frac{Q'}{T'} > 0.$$

Ne consegue che un processo di questo genere è escluso da entrambe le formulazioni. Analogamente, se una quantità di calore Q' venisse completamente trasformata in lavoro, varrebbe la relazione $\oint(dQ/T)=Q'/T>0$, che è esclusa dalla disuguaglianza di Clausius.

La seconda legge può essere espressa servendosi dell'entropia. La disuguaglianza di Clausius non dà alcuna informazione diretta sulla variazione di entropia in un ciclo irreversibile, poiché non esiste alcun collegamento immediato tra il calore assorbito e la variazione di entropia, a meno che il ciclo sia reversibile. Si può però dedurre dalla disuguaglianza che in qualsiasi processo irreversibile l'entropia aumenta. A questo fine si prenda in esame un sistema chiuso che subisca una variazione *spontanea* (ossia irreversibile) dallo stato A allo stato B. Si supponga altresí che la trasformazione spontanea possa essere seguita da un processo *reversibile* che riconduca il sistema allo stato A (questo stadio implicherebbe necessariamente un'interazione con l'ambiente circostante). Globalmente si ha un ciclo irreversibile per cui deve sussistere $\oint(dQ/T)<0$. Si può suddividere l'integrale, ottenendo $\int_A^B(dQ/T)-\int_B^A(dQ/T)<0$, ove il primo integrale appartiene alla trasformazione spontanea e il secondo a quella reversibile. Questo secondo integrale è per definizione uguale a (S_A-S_B) (S_A

e S_B sono rispettivamente le entropie degli stati A e B) e quindi si può scrivere la disuguaglianza:

$$S_B - S_A > \int_A^B \frac{dQ}{T}.$$

Ma la trasformazione spontanea ha luogo in un sistema chiuso e di conseguenza durante il processo non si ha assorbimento di calore. Ne discende che $S_B-S_A>0$, o $S_B>S_A$, il che significa che l'entropia è *aumentata* dal valore S_A a un valore piú elevato S_B.

È possibile allora riformulare la seconda legge della termodinamica: l'entropia di un sistema chiuso può soltanto crescere, o, nelle parole di Clausius, l'entropia di un sistema chiuso tende a un massimo.

Si possono cosí sintetizzare gli sviluppi sinora descritti: la legge di conservazione dell'energia esprime l'equivalenza del calore e del lavoro come forme di energia reciprocamente trasformabili, e esclude un perpetuum mobile del primo ordine. Tuttavia l'analisi di Carnot dimostra che non tutte le trasformazioni possibili secondo questa legge possono veramente accadere in natura. Un perpetuum mobile del secondo ordine è *consentito* dalla legge di conservazione dell'energia (la prima legge della termodinamica), ma *escluso dalla seconda legge*. Infine si è visto come Clausius abbia dimostrato che l'esclusione di un perpetuum mobile del secondo ordine deriva dal principio generale dell'aumento dell'entropia nei sistemi chiusi.

È illuminante esaminare la connessione tra la prima e la seconda legge in relazione alla disponibilità di energia. La prima legge si riferisce soltanto alla «contabilità» generale dell'energia di un sistema. La seconda legge ha invece a che fare con l'utilità dell'energia, cioè con la sua disponibilità a compiere lavoro. Il fatto che l'entropia di un sistema chiuso aumenti sempre significa che i processi spontanei in un sistema chiuso tenderanno a trasformare in energia termica le quantità eccedenti di energia meccanica, elettrica o chimica. Poiché non è possibile convertire completamente l'energia termica in lavoro, il significato di questo processo è che, se si lascia a se stesso un sistema chiuso, una quantità di energia via via minore sarà utilizzabile per ottenere lavoro. Si afferma talora che in natura si riscontra una tendenza verso la *degradazione dell'energia*, ossia che l'energia tende a trasformarsi in energia inutile. Piú un sistema è prossimo allo stato di massima entropia, piú la sua energia è degradata.

A conclusione del paragrafo, un'ultima osservazione: il principio di aumento dell'entropia nella termodinamica classica è un postulato (naturalmente ben confermato). Esso non spiega perché i fenomeni naturali si orientino in una certa direzione, limitandosi a esprimere il fatto in modo elegante. Nel paragrafo successivo si dimostrerà che la teoria della meccanica statistica ha contribuito a una migliore comprensione del concetto di entropia, conferendo cosí al postulato un miglior potere esplicativo.

2. *L'entropia nella meccanica statistica.*

2.1. Necessità di una teoria meccanicistica.

La termodinamica classica, nonostante il suo successo nel «salvare i fenomeni», aveva un serio inconveniente dal punto di vista filosofico. A partire dal XVII secolo, molti fra i piú grandi scienziati difesero un atteggiamento filosofico che prese il nome di *meccanicismo*. Secondo l'ottica meccanicistica i movimenti delle particelle materiali dovrebbero essere l'unico fondamento per spiegare i fenomeni naturali. Il meccanicismo si proponeva di eliminare, dal regno della fisica, le «essenze» o «sostanze» diverse dalla materia e dal moto. Il concetto di entropia, qual è introdotto nella termodinamica classica, sembrava irriducibile alla terminologia meccanicistica e quindi pareva manifestare un carattere un po' occulto, inadatto alla scienza moderna. Inoltre, il successo dell'atomismo tanto in fisica quanto in chimica apportò una nuova sfaccettatura alla visione meccanicistica: i fenomeni al livello sperimentale – il macrolivello – dovrebbero essere spiegati in base ai processi fisici che hanno luogo al livello atomico – il microlivello. La termodinamica classica era limitata al macrolivello ed era incapace di assicurare l'agognato legame tra mondo osservabile e mondo degli atomi. Il grande risultato della meccanica statistica fu la determinazione dell'anello mancante, che permise di trasformare il principio di aumento dell'entropia, apparentemente refrattario a un'interpretazione meccanicistica, in uno dei successi piú clamorosi di questa filosofia.

2.2. L'approccio statistico.

Le leggi classiche del moto possono pienamente descrivere il movimento di un sistema costituito da un piccolo numero di particelle, a patto che siano date le condizioni iniziali (ad esempio le posizioni e i momenti iniziali delle particelle). Accingendosi a descrivere il comportamento di un macrosistema quale risulta dai movimenti delle particelle che lo compongono, ci si trova di fronte a un problema enorme: un macrosistema può essere formato da un gran numero di atomi o molecole: una mole di gas contiene ad esempio $0{,}6 \times 10^{24}$ molecole. È impossibile determinare l'esatto stato di un sistema di questo genere, cioè determinare le condizioni di tutte le sue particelle in ogni istante considerato. È però possibile sviluppare una teoria *statistica* del comportamento di tali grandi sistemi, e dimostrare che le funzioni caratterizzanti uno stato di un sistema macroscopico, quali la temperatura, la pressione, l'entropia, ecc., sono esprimibili in termini di *medie* relative alle funzioni che caratterizzano stati di particelle individuali. Per l'approccio statistico, l'elevato numero di particelle al livello microscopico è naturalmente un vantaggio anziché una difficoltà. Come in qualsiasi altra teoria statistica, le fluttuazioni dalle distribuzioni medie sono tanto minori quanto piú ampio è il campione.

Secondo la teoria statistica ogni macrostato si può visualizzare come descrizione sintetica; in altre parole, ogni macrostato, dato dai valori delle variabili macroscopiche, può essere la conseguenza di moltissime diverse configurazioni a livello atomico. Un macrostato corrisponde cosí a un gran numero di microstati diversi soltanto al livello atomico e indistinguibili al macrolivello.

Le basi di questa nuova teoria statistica, nota come meccanica statistica, furono gettate da James Clerk Maxwell, Ludwig Boltzmann [1896-98] e da Gibbs [1902]. Poiché la teoria di Maxwell si occupa soprattutto della distribuzione dell'energia, e non dell'entropia, non la si tratterà in questa sede.

2.3. Probabilità e disordine.

Macrostati diversi si differenziano per il numero di modi in cui possono essere composti da microparticelle. Laddove alcuni macrostati possono essere realizzati mediante configurazioni diverse delle particelle costituenti, altri «sintetizzano» soltanto un numero limitato di disposizioni possibili. Maggiore è il numero delle disposizioni possibili che dànno un determinato stato, piú elevata è la probabilità di trovare il sistema in quello stato particolare.

Per illustrare ciò, si considerino le possibili disposizioni di un sistema costituito da cinquanta palle da biliardo rosse e cinquanta bianche. Agitando il contenitore delle palle si possono ottenere varie disposizioni. Il numero totale delle configurazioni possibili è $100! = 100 \times 99 \times 98 \times \ldots \times 2 \times 1$. Poiché non c'è alcun motivo di ritenere che una certa configurazione sia piú probabile di un'altra, si attribuisce a ognuna delle configurazioni possibili la stessa probabilità: $1/100!$ Di fatto, se le palle non sono numerate o in qualche altro modo contrassegnate, è possibile scambiare una qualunque delle palle rosse o delle palle bianche senza ottenere una configurazione distinguibile dalla precedente. Il numero totale di configurazioni *distinguibili* è quindi «soltanto» $100!/(50! \times 50!) = 10^{29}$. Si prenda ora in esame la probabilità di una distribuzione estremamente regolare delle palle, ad esempio una distribuzione tridimensionale a scacchiera, ove ogni palla rossa ha vicine soltanto palle bianche. Tra tutte le 10^{29} possibili configurazioni, soltanto due avranno come risultato questa distribuzione regolare, la cui probabilità è quindi soltanto $2/10^{29}$. Una distribuzione regolare con un «difetto» si può realizzare in duemilacinquecento modi (ogni palla bianca si può scambiare con ognuna delle palle rosse); la probabilità di una configurazione con un difetto è quindi già duemilacinquecento volte superiore a quella della distribuzione regolare. Procedendo con l'analisi lungo le stesse direttrici ci si accorge rapidamente che la probabilità di una distribuzione è tanto maggiore quanto meno essa è regolare. Non deve allora stupire che agitando il contenitore in pratica non si raggiunga mai la distribuzione a scacchiera e si osservino soltanto disposizioni disordinate, nonostante il fatto che la probabilità a priori delle configurazioni corrispondenti alla distribuzione a scacchiera non sia inferiore a quella di qualsiasi altra configurazione.

Al posto delle cento palle da biliardo, si consideri un sistema costituito da due gas diversi, in uguali condizioni di temperatura e pressione, separati da una parete. Eliminando la parete, si verifica una miscelazione dei gas sino

al raggiungimento di una distribuzione uniforme. Con ragionamenti analoghi a quelli applicati alle palle da biliardo si può facilmente comprendere che il processo di miscelazione è in effetti il fenomeno piú probabile. Vi è soltanto una differenza: poiché il numero delle molecole componenti il sistema è molto superiore al numero delle palle da biliardo nell'esempio, la probabilità di una distribuzione regolare è talmente piú bassa da poter essere di fatto trascurata. La transizione del sistema dallo stato iniziale dei due gas separati allo stato di miscelazione si può descrivere come transizione a uno stato di casualità o di disordine. La si può anche descrivere in termini probabilistici come transizione a uno stato di probabilità superiore.

2.4. Probabilità ed entropia.

Secondo la termodinamica classica, la miscelazione di due gas del tipo considerato si associa a un aumento di entropia. Gibbs ha dimostrato che l'entropia di una miscela di gas perfetti è pari alla somma delle entropie che ogni gas possiederebbe separatamente se occupasse da solo l'intero volume della miscela. Si potrebbe allora affermare che il processo descritto *in senso classico* come capace di dare un aumento di entropia, può essere rappresentato *statisticamente* come transizione a una distribuzione piú probabile. Questo esempio suggerisce la possibilità di stabilire una relazione generale tra l'entropia e la probabilità di uno stato termodinamico. Per trovarne il collegamento matematico si deve osservare che l'entropia del sistema è una funzione *additiva*: se si immagina che due sistemi separati A e B siano un sistema unico, l'entropia del sistema combinato S_{AB} è data dalla *somma* delle entropie, S_A+S_B. D'altra parte, il numero di stati possibili nel sistema combinato è il *prodotto* del numero degli stati possibili nel sistema A per il numero degli stati possibili nel sistema B. Ne consegue che l'entropia e il numero delle configurazioni possibili devono essere legati da una funzione logaritmica.

La discussione precedente, che ha prospettato la possibilità di una teoria statistica dell'entropia, naturalmente è molto semplificata. Considerazioni essenzialmente simili, ma molto piú profonde nei particolari, condussero Boltzmann a proporre la famosa equazione, che costituisce la definizione statistica di entropia: $S=k \lg m$; ove S è l'entropia di un sistema in un dato stato termodinamico; m è il numero di configurazioni nelle quali si può realizzare lo stato termodinamico; k è una costante, nota come costante di Boltzmann, con valore pari a $1{,}38\times 10^{16}$ erg/grado. Questa costante esprime il legame tra la teoria classica e quella statistica. La si ottiene in base alla formula $k=R/N_0$, in cui R è la costante dei gas (il cui valore dipende dalle altre unità usate) e N_0 è il numero di Avogadro, il numero di atomi o molecole per mole, approssimativamente pari a $0{,}6\times 10^{24}$.

(Volendo mantenersi fedeli alla verità storica, si dovrebbe far rilevare che questa forma della definizione è una variazione apportata da Max Planck a quella di Boltzmann).

La definizione viene generalizzata nel modo seguente: anziché considerare *un* sistema e calcolarne il numero delle disposizioni possibili, si supponga di avere a disposizione un ipotetico *insieme* di copie dello stesso sistema. Ogni sistema dell'insieme si troverà in uno dei possibili microstati e l'insieme, globalmente, è definito quando è nota la frazione dei suoi sistemi che si trova in ciascun microstato. Se nel microstato i-mo si ha una frazione P_i dei sistemi, l'entropia S è definita

$$S=-K\sum_i P_i \lg P_i$$

ove K è la costante di Boltzmann. Se l'insieme è distribuito su m microstati della stessa probabilità, allora tutte le P_i hanno lo stesso valore: $P_i=1/m$.

In questo caso si ottiene

$$S=-K\sum_i \frac{1}{m}\lg\frac{1}{m}=-Km\frac{1}{m}\lg\frac{1}{m}=K\lg m$$

che coincide con la definizione precedente. La diversità tra la presentazione introdotta da Gibbs, basata sugli insiemi, e il precedente approccio di Boltzmann verrà spiegata in seguito.

L'esempio del processo di miscelazione non giustifica naturalmente la definizione statistica di entropia in generale, ma si limita a renderla plausibile in un caso particolare. Comunque, ne deriva che il concetto statistico di entropia dà risultati identici al concetto classico in tutti i casi in cui si applica quest'ultimo, valendo inoltre anche per quei casi per cui il concetto classico è inutilizzabile o decisamente poco conveniente.

Nel paragrafo precedente si è chiarito che il legame tra entropia e passaggio di calore è valido soltanto per processi *reversibili*. Non si può effettuare alcun calcolo diretto della variazione di entropia in un processo irreversibile. D'altro lato la definizione statistica non è limitata a un tipo particolare di processo. Si è anche sottolineato che il calcolo dei valori *assoluti* dell'entropia, e non soltanto delle variazioni di entropia, richiede l'attribuzione di un'entropia nulla ai cristalli ideali allo zero assoluto. La definizione statistica rende meno arbitraria questa assegnazione. Allo zero assoluto si suppone che un cristallo ideale sia completamente ordinato, ovvero che il numero delle configurazioni possibili sia uno, con conseguente entropia nulla.

I vantaggi formali e pratici della definizione statistica di entropia sono soltanto secondari a confronto del suo piú grande merito, cioè la chiarezza concettuale. Il concetto di entropia, come abbiamo spiegato, fu introdotto da Clausius per esprimere la direzionalità dei fenomeni naturali. Nonostante la chiarezza dell'analisi di Clausius, rimase qualche nebulosità sull'entità da lui definita. La nebbia è oggi dissipata. Il principio di aumento dell'entropia ha ora questa forma: Gli stati di entropia elevata sono piú probabili di quelli con entropia piú bassa. Le transizioni osservate a stati di entropia maggiore sono transizioni a stati di probabilità maggiore. Il profumo si diffonde nella stanza non a causa della crescita di qualche misteriosa entità irriducibile, ma semplicemente grazie al fatto che una sua distribuzione omogenea per l'intera stanza è molto piú probabile di qualsiasi altra distribuzione.

La legge classica di aumento dell'entropia è una legge di carattere *teleologico*: interpreta la direzionalità dei fenomeni come effetto di una *tendenza* in natura *verso* l'aumento di entropia. Il tipo di causalità implicato in questa interpretazione dei fenomeni è intenzionale, o finalizzato. La filosofia meccanicistica cerca di abolire questo genere di causalità dalla scienza e, in effetti, il concetto statistico di entropia è privo di caratteri teleologici. Un sistema fisico non «preferisce» stati di entropia piú elevata; non si riscontra in natura alcuna «tendenza» a degradare l'energia: l'apparente tendenza che si verifica a livello macroscopico si può spiegare in termini puramente meccanici e statistici.

Ovviamente il principio statistico di aumento dell'entropia differisce nel suo contenuto da quello classico. Per il fatto stesso di essere un principio statistico, esso ammette delle deviazioni. Secondo il principio classico l'entropia deve aumentare e ogni sua diminuzione osservata contraddirebbe la legge. Tuttavia la versione statistica della legge predice soltanto un aumento medio di entropia. Un sistema chiuso lasciato a sé si avvicinerà *in media* alla distribuzione piú probabile e fluttuerà nei suoi pressi, ma possono verificarsi occasionali transizioni a uno stato meno probabile, e se ne può anzi dimostrare l'inevitabilità.

È difficile sopravvalutare il significato di questa differenza tra la legge classica e la legge statistica. Errate interpretazioni del carattere della legge statistica hanno partorito generazioni di paradossi che parvero a lungo minacciare la coerenza della meccanica statistica. Prima di discutere alcuni di questi paradossi occorre però aggiungere un altro rilievo sui fondamenti della meccanica statistica.

2.5. Presentazione di Boltzmann vs presentazione di Gibbs.

Sviluppando la sua teoria statistica, Boltzmann aveva in mente l'evoluzione nel tempo di un determinato sistema termodinamico. La validità dell'analisi statistica si basa sul fatto che il sistema è formato da un gran numero di molecole, ovvero che il suo insieme statistico è costituito dalle molecole che compongono il sistema in esame. L'analisi si applica soprattutto ai gas e si devono fare diverse ipotesi sui tipi d'interazione esistenti tra le molecole. Nella terminologia boltzmanniana il principio di aumento dell'entropia significa che un sistema non in equilibrio si evolverà nel tempo verso uno stato piú probabile e raggiungerà l'equilibrio fluttuando attorno allo stato piú probabile. Si potrebbe anche dire che, nel modello di Boltzmann, il sistema *trascorrerà piú tempo nello stato piú probabile* (o in un suo intorno) che in qualsiasi altro stato.

Nella sua formulazione della meccanica statistica, Gibbs si appellò a un modello diverso. Il suo insieme non è formato dalle particelle che compongono il sistema *reale* in esame, ma è piuttosto una raccolta *ipotetica* di sistemi identici, tra i quali quello reale è considerato soltanto un esponente. Lo stato piú probabile sarà occupato dalla *piú alta percentuale* di sistemi, e quindi è assai probabile che il nostro sistema reale appartenga a quella frazione. Si capisce immediatamente che la validità dell'analisi statistica nella formulazione di Gibbs

non è influenzata né dal numero delle particelle componenti il sistema, né dalle loro interazioni. Il suo linguaggio degli insiemi è quindi adeguato non soltanto per i gas, nei quali l'interazione tra le particelle è debole, ma anche per i solidi, i liquidi o qualsiasi altro sistema quantico. Di conseguenza, i fisici si servono oggi in generale piú del modello di Gibbs che di quello di Boltzmann. Praticamente tra le due teorie non sussiste differenza alcuna: esse portano agli stessi risultati; la diversità esiste soltanto in relazione alla giustificazione della teoria, e non alla teoria stessa.

2.6. Paradossi di irreversibilità.

Nel 1876 Joseph Loschmidt mosse un'obiezione alla teoria di Boltzmann: la teoria statistica di Boltzmann è un tentativo di ridurre la termodinamica dei macrosistemi alla meccanica delle microparticelle. Ma le leggi della meccanica classica sono invarianti per l'inversione del tempo e quindi tutti i processi da loro controllati devono essere reversibili. Prendendo ancora in esame una massa di gas in un sistema chiuso, in evoluzione verso l'equilibrio termodinamico, è possibile, secondo l'argomentazione di Loschmidt, che ad un certo momento tutte le velocità delle molecole invertano la direzione (il che equivale all'inversione del tempo). A partire da quell'istante il sistema ribalterà la sua evoluzione e, di conseguenza, diminuirà di entropia, in contraddizione con la seconda legge della termodinamica. Non è soltanto impossibile *ridurre* la termodinamica alla meccanica statistica: le due teorie, secondo Loschmidt, sono *incompatibili*.

Un altro paradosso fu scoperto verso la fine del secolo da Ernst Zermelo. Egli si appoggiò a un teorema, dimostrato da Poincaré, secondo cui ogni sistema meccanico limitato e conservativo si evolverà con un moto quasi-periodico, in modo da ritornare sempre (se gli si concede un tempo sufficiente) a un qualsiasi stato iniziale dato. Zermelo sostiene che da questo teorema discende che un sistema non può evolversi indefinitamente nella stessa direzione e che *certamente* avverranno transizioni a stati caratterizzati da una minore probabilità (ossia transizioni a stati di entropia minore). Anche questa conclusione sembra smentire la seconda legge della termodinamica.

Entrambi i paradossi scaturiscono da una interpretazione lievemente erronea della legge statistica. Si deve sempre tenere presente la distinzione tra macrolivello e microlivello. Ogni microstato è di fatto realizzabile, e lo è un numero infinito di volte in base al teorema di Poincaré; ma nell'insieme prevarranno per un tempo sufficiente all'osservazione soltanto quei macrostati che rappresentano la frazione piú ampia di microstati. Cosí, la tendenza generale *osservabile* dell'intero sistema consisterà nell'approssimarsi allo stato di equilibrio o di massima entropia. In effetti si verificano fluttuazioni rispetto a quello stato, ma quanto piú sono ampie, tanto meno sono probabili e ci lasciano la possibilità di vederle accadere. Il paradosso di Zermelo costituisce in realtà per le idee di Boltzmann una sfida piú profonda di quanto lasci pensare la spiegazione fornita: lo si riesaminerà nell'ultimo paragrafo.

2.7. Computazioni effettive dell'entropia statistica.

Come si è visto, nella meccanica statistica l'entropia è una misura del numero delle possibili configurazioni di un sistema fisico. Resta aperto il problema della determinazione pratica di questo numero. Il compito di accertarlo nei vari casi è il problema principale della meccanica statistica. Quando un sistema è caratterizzato dalla sua energia, la questione si riduce alle modalità di distribuzione dell'energia tra i vari costituenti. Di solito il problema è rovesciato: la meccanica quantistica ci fornisce i possibili livelli energetici del sistema, o i suoi possibili stati quantici, lasciandoci il compito di stabilire come siano distribuiti tra i vari stati quantici i costituenti del sistema (molecole, atomi, elettroni, ecc.). In certi casi (ad esempio in un cristallo solido a temperatura ambiente) soltanto un numero esiguo dei livelli energetici è occupato dalle particelle; vi sono cioè piú particelle che livelli praticamente disponibili. In altri casi (ad esempio un gas in condizioni normali) il numero delle particelle è molto inferiore a quello dei livelli disponibili. È impossibile addentrarci in questa sede nei dettagli del calcolo, ma nondimeno è necessario illustrare un'importante considerazione, essenziale ai nostri scopi.

Un esempio semplice: distribuire due ricompense fra tre bambini. Il numero delle possibilità di farlo dipenderà *a*) dalla natura delle ricompense, e *b*) dalle condizioni in cui le ricompense sono date. Se le ricompense sono distinguibili (poniamo due libri diversi) e si consente che un bambino ne vinca due, è possibile distribuirle in nove modi diversi. Se le ricompense sono tali che l'interscambio tra due bambini non produca una nuova possibilità (si supponga che ogni ricompensa sia un biglietto da mille lire), le possibilità sono soltanto sei. Se poi s'impone la restrizione di una vincita limitata a una ricompensa per bambino, si resta con tre sole possibilità.

S'immagini ora che i bambini rappresentino gli stati quantici, e che le ricompense rappresentino le particelle. È chiaro che per elencare le possibili distribuzioni occorre sapere *a*) se si possono considerare le particelle distinguibili l'una dall'altra, e *b*) se esiste una limitazione riguardo al numero delle particelle in grado di occupare uno stato quantico. I primi calcoli effettuati da Maxwell e Boltzmann, con particelle distinguibili e senza restrizioni relative alla distribuzione tra gli stati, diedero un numero di possibili configurazioni che ben presto si dimostrò troppo elevato. Si dovette ammettere che le microparticelle costituenti un sistema fisico sono *indistinguibili* l'una dall'altra, per cosí dire prive d'individualità, e che il numero dei diversi microstati producenti un determinato macrostato è invero molto inferiore a quello stimato da Maxwell e Boltzmann. Il calcolo statistico fondato sull'indistinguibilità delle particelle fu sviluppato indipendentemente da Bose e Einstein ed è perciò noto come statistica di Bose-Einstein. Ci si è poi accorti che anche la statistica di Bose-Einstein non è di applicazione generale. Esistono sistemi (ad esempio gli elettroni in un atomo) che obbediscono al principio di esclusione di Pauli, ovvero al principio secondo cui ogni stato quantico può essere occupato *al massimo* da una particella. Il numero delle distribuzioni possibili, e quindi anche l'entropia di tali sistemi, viene cosí ulteriormente ridotto. La statistica che descrive questi sistemi è detta (dal nome dei fisici che la misero a punto) statistica di Fermi-Dirac. Le particelle sono caratterizzate dal tipo di statistica cui sono sottoposte e si dividono in fermioni, che obbediscono alla statistica di Fermi-Dirac, e bosoni, che seguono la statistica di Bose-Einstein. Quando il numero dei livelli energetici utilizzabili è molto superiore al numero delle particelle, è fortemente improbabile che un livello venga occupato da piú di una particella; in questo caso la statistica di Bose-Einstein fornirà risultati identici a quella di Fermi-Dirac. In altri casi la differenza tra le due è molto significativa per il calcolo dell'entropia e di altre funzioni termodinamiche.

2.8. Paradosso di Gibbs.

La distinzione tra statistica di Bose-Einstein e trattamento classico in base alla statistica di Maxwell-Boltzmann aiuta a risolvere un paradosso noto come paradosso di Gibbs. Si prenda nuovamente in esame la miscelazione di due gas, ognuno dei quali occupi all'inizio un volume v. Dopo la miscelazione i due gas combinati occupano il volume $2v$. Come si è visto, l'entropia cresce nel processo di mescolamento; laddove prima della miscelazione l'entropia dei due gas era data dalla somma delle rispettive entropie, S_1+S_2, nello stato misto si può dimostrare che l'entropia è cresciuta al valore $S_1+S_2+2kN\log 2$, ove N è il numero delle molecole di ciascun gas. Questi risultati sembrano paradossali in quanto, nelle considerazioni che li hanno determinati, nulla dipende del fatto di mescolare due gas *diversi*. Sembra che lo stesso risultato si ottenga anche miscelando due volumi del *medesimo* gas. Tuttavia l'eliminazione di una suddivisione in un gas omogeneo non pare costituire un cambiamento fisico significativo, e di conseguenza in questo processo l'entropia non dovrebbe aumentare. Il problema si risolve trattandolo in base alla statistica di Bose-Einstein: dati due gas diversi, le loro molecole sono *distinguibili* e quindi, scambiando due particelle dei due gas diversi, si ottiene un nuovo stato; se, d'altra parte, si mescolano due volumi del medesimo gas, lo stesso scambio non produrrà uno stato diverso e il numero delle configurazioni possibili sarà *minore* rispetto al caso dei due gas. Si può dimostrare che il numero delle configurazioni si riduce dello stesso fattore che costituiva il problema: ossia di $2kN\log 2$.

2.9. Termodinamica dei sistemi aperti.

Sinora ci si è occupati del concetto di entropia nella termodinamica classica e statistica. In entrambi i casi si è soprattutto trattato di sistemi *chiusi*, che non interagiscono con l'ambiente e possono approssimare uno stato di equilibrio termodinamico. Tuttavia, molti dei sistemi nei quali ci s'imbatte in natura, e per i quali c'è interesse, sono, di fatto, sistemi *aperti*, caratterizzati da uno scambio costante di materia e energia con l'esterno. Lo studio della ter-

modinamica di tali sistemi è della massima importanza per la comprensione dei sistemi *biologici*. Sono stati compiuti tentativi per generalizzare i risultati della termodinamica in modo da abbracciare anche il comportamento dei sistemi aperti, ma, sino ad oggi, le nuove teorie sono largamente incomplete. Una loro discussione va al di là degli obiettivi del nostro articolo: se ne può trovare un riferimento nella bibliografia.

3. *L'entropia nella teoria dell'informazione.*

Come si è osservato, l'entropia nella meccanica statistica è una misura del disordine o casualità di un sistema fisico. Uno stato di maggiore entropia è uno stato realizzabile in un maggior numero di modi. La funzione statistica entropia dipende esclusivamente dalla distribuzione di probabilità e soltanto la costante moltiplicativa – costante di Boltzmann – dipende dalle funzioni termodinamiche: viene cosí fornito il legame tra termodinamica statistica e classica. Il passo successivo consiste nell'affermare che il concetto di entropia è pienamente separabile da qualunque contesto fisico e che ad ogni distribuzione di probabilità si può associare una funzione entropia. Quest'entropia astratta non verrà naturalmente misurata in unità termodinamiche, ma sarà un numero puro. Verranno ora discussi il significato e le applicazioni dell'entropia statistica astratta, e la relazione tra questo concetto e quello termodinamico.

3.1. Informazione.

La teoria dell'informazione s'interessa dello studio della trasmissione di segnali o messaggi da un meccanismo (esseri umani inclusi) a un altro. La gamma dei problemi a cui si può applicare la teoria è assai ampia; linguaggio, letteratura, telecomunicazioni, musica, meccanismo genetico degli organismi viventi, sono tutti esempi di fenomeni la cui analisi rientra nella teoria dell'informazione. La teoria dell'informazione si occupa principalmente non del *significato* dei messaggi, ma della loro *struttura statistica*. La teoria indaga le condizioni che porterebbero a una migliore o piú efficiente trasmissione dell'informazione, definita in termini formali statistici e non in termini semantici o estetici. Tuttavia, come possono attestare alcuni degli esempi discussi, struttura e significato non sono completamente indipendenti.

I fondamenti della teoria matematica dell'informazione furono gettati da Shannon nel 1948 [cfr. Shannon e Weaver 1949]. La sua analisi si fonda sulle seguenti considerazioni: il *reale* messaggio che si trasmette è un messaggio scelto tra un insieme di messaggi *possibili*. Quanto piú elevato è il numero dei messaggi possibili, tanto maggiore è la quantità d'informazione prodotta al momento della scelta effettiva di un messaggio. Il numero dei messaggi possibili può quindi provvedere, secondo Shannon, *una misura dell'informazione* determinata da una particolare scelta.

Analogamente, nell'esame di un problema implicante un certo numero di

risposte possibili, un'ulteriore informazione sulla situazione ridurrà il numero delle possibilità. Se l'informazione fosse completa, si potrebbe restare con un'unica possibilità. In altre parole, l'incertezza iniziale relativa al problema lascia il passo, sulla base dell'informazione pertinente, alla certezza o predicibilità. La quantità d'informazione prodotta (per esempio, grazie a un certo esperimento o segnale, ecc.) può per contro essere definita in funzione del numero delle risposte possibili con e senza l'informazione. In entrambi i casi, *piú informazione* significa *eliminazione di piú incertezza*.

Si supponga di avere uno schema di n eventi possibili e si ipotizzi di conoscere le probabilità che questi eventi si verifichino, $P_1, P_2, \ldots, P_n$. S'immagini pure che ciò costituisca tutto quello che si sa sul problema. Shannon è riuscito a dare una misura matematica dell'incertezza insita nel problema, ovvero dell'informazione ricavata grazie alla sua eliminazione. La funzione che egli ha dimostrato utile come misura è

$$-K\sum_{i=1}^{n} P_i \lg P_i$$

ove K è una costante che determina l'unità d'informazione.

Varie sono le considerazioni che rendono plausibile questa definizione. Ad esempio, se tutti gli eventi hanno probabilità 0, salvo uno con probabilità 1 ($P_j = 1$, $P_i = 0$ per $i \neq j$), si sarebbe tentati di affermare che non v'è affatto incertezza. Come si può osservare, la funzione definita soddisfa a questo requisito e dà come risultato zero. È altresí chiaro che lo schema col grado massimo d'incertezza è quello in cui tutti gli eventi sono ugualmente probabili, $P_i = 1/n$: è ciò che si verifica con la misura di Shannon. Si possono fare altre considerazioni, omesse in questa sede. Il principale successo della definizione di Shannon sta nel fatto che egli è riuscito a dimostrare che la sua funzione non è soltanto una misura possibile, ma è l'*unica* funzione (a meno di una costante) che soddisfa i requisiti matematici.

3.2. Informazione ed entropia.

A questo punto il lettore deve aver compreso che la funzione di Shannon è completamente analoga (ancora a meno di una costante) alla definizione statistica di entropia. In effetti, grazie a questa analogia matematica, la funzione di Shannon è detta *entropia* dello schema. Piú è alta l'entropia di un sistema di eventi, maggiore è l'informazione associata ad una scelta particolare di un evento specifico. Anche questo è intuitivamente evidente: un risultato scelto tra dieci possibili toglie piú incertezza e quindi porta con sé piú informazione di un risultato scelto tra due sole possibilità. Shannon si è servito della stessa funzione per misurare sia l'entropia sia l'informazione. Altri autori definiscono l'informazione come entropia negativa (negentropia), adducendo a motivo che l'informazione si dovrebbe associare all'incertezza *soppressa*. La differenza sta soltanto nel segno, non nel valore delle funzioni. L'entropia qui definita può naturalmente essere associata a qualunque funzione di distribuzione e non è

misurata in unità termodinamiche. Prima di prendere in esame la relazione tra il concetto astratto e il concetto fisico di entropia, s'impone in via preliminare una discussione della significatività del concetto quale è usato nella teoria dell'informazione.

3.3. Entropia e ridondanza.

I messaggi sono trasmessi in un linguaggio di simboli o segnali. I segnali possono essere parole di un linguaggio naturale, lettere di un alfabeto, note musicali, punti e linee come in un messaggio telegrafico, oppure una serie di composti chimici sul tipo di quelli che trasmettono le informazioni genetiche tramite il codice genetico. Quale che sia il sistema di segnali, la quantità d'informazione convogliabile dipende dal numero delle possibili disposizioni dei segnali. Come si è notato, la quantità d'informazione sarà massima quando tutti i segnali sono ugualmente probabili e nessun vincolo è imposto sulla loro relativa dipendenza. Ogni introduzione di un qualsiasi vincolo esclude immediatamente alcune delle disposizioni possibili, riducendo cosí la quantità d'informazione potenziale. Le regole ortografiche di una lingua naturale costituiscono un esempio di tali vincoli. Se, ad esempio, ogni parola inglese dovesse includere una vocale (o la lettera *y*), questa regola ridurrebbe la libertà di scelta nella costruzione delle parole. Allo stesso modo si trovano «regole ortografiche» nella musica, nel balletto o nell'informazione genetica. Ogni sistema sul quale gravano vincoli o «regole ortografiche» possiede un'*entropia minore* di quella che avrebbe in assenza delle regole. In altre parole, *struttura* e *entropia* sono due caratteri esprimenti aspetti contraddittori, e ogni linguaggio deve equilibrare questi aspetti in modo da raggiungere un'elevata entropia, o un alto contenuto informazionale, da un lato, e una buona struttura dall'altro.

Si può anche introdurre una misura matematica del grado di struttura: si calcola l'entropia di una determinata informazione-sorgente (con vincoli) e la si confronta con il valore massimo che l'entropia avrebbe potuto avere in assenza di vincoli. Il rapporto tra l'entropia reale e quella massima si chiama *entropia relativa* della sorgente. La quantità (1 — entropia relativa) esprime la frazione della composizione del messaggio che è determinata non dalla *libera scelta* della sorgente, ma dalle *regole di struttura* del linguaggio usato. Questa quantità è detta *ridondanza* del messaggio. Il termine 'ridondanza' è stato scelto per il fatto che, in un certo senso, i vincoli o le regole ortografiche di un linguaggio rendono ridondante una parte dei segnali trasmessi, la parte che avrebbe potuto essere indovinata o completata sulla base della conoscenza delle regole ortografiche. È assai probabile che una parola inglese terminante con le lettere *-tio-* presenti come lettera finale un'addizionale *-n*. Se quest'ultima manca in un telegramma, è facile introdurvela. In un certo senso essa è quindi ridondante.

L'analisi dei linguaggi in termini della loro entropia e ridondanza è interessante di per sé, ma offre anche implicazioni pratiche di vasta portata. Ad esempio, è necessaria in relazione a problemi di *codificazione*. Nel costruire un codice, due sono gli obiettivi: da un lato si vuole che il codice sia il piú conciso possibile (bassa ridondanza ed elevata entropia); dall'altro si desidera essere in grado di rintracciare e correggere ogni errore eventualmente insinuatosi durante la trasmissione. Il modo piú semplice e meno sofisticato per assicurare l'affidabilità consiste nel ripetere ogni messaggio, ma, tenendo presente l'analisi precedente, esiste un meccanismo migliore: aumentare la ridondanza del codice. Come nel caso del telegramma, i vincoli imposti favoriranno la correzione degli errori. In ogni processo di trasmissione ci si trova di fronte a disturbi o rumori, ovvero a informazioni provenienti da fonti esterne. La ridondanza aiuta a combattere il rumore. Un buon codice è quindi un codice nel quale si è realizzato il giusto dosaggio tra l'entropia, che garantisce la *varietà*, e la ridondanza, che assicura l'*affidabilità*. Si è calcolato che la ridondanza della lingua inglese è attorno al 50 per cento, il che significa che metà delle lettere e parole impiegate sono determinate dalle regole di struttura del linguaggio e non dalla libera scelta di chi parla o scrive. Ecco ora altri due esempi scelti nell'ampio numero di fenomeni cui si può applicare questo tipo di analisi.

Entropia e musica. Hiller [1959] sostiene che talvolta la musica è definita come compromesso tra caos e monotonia. È difficile che un suono ripetuto a intervalli uguali sia considerato un brano musicale e, d'altra parte, anche una sequenza casuale di suoni, irrispettosa di regole di struttura, non offre molto all'ascoltatore. Nei termini dell'analisi precedente, si può affermare che la musica, al pari del linguaggio, tenta di realizzare un equilibrio tra entropia e ridondanza. Con quest'affermazione non si vogliono ridurre gli aspetti estetici della musica alla sua struttura formale, ma soltanto riconoscere che certamente la struttura formale non è estranea al piú profondo livello estetico.

L'analisi della musica secondo la teoria dell'informazione è d'altronde significativa sotto due aspetti. In primo luogo essa possiede un valore teorico, consentendo di definire e riconoscere stili diversi di musica in base alla loro struttura statistica; secondariamente, dopo aver compiuto l'analisi degli stili, è possibile servirsene per comporre musica in uno stile particolare o allo scopo d'inventare nuove regole strutturali. Nell'articolo di Hiller si possono rinvenire esempi di entrambi gli aspetti dell'analisi. A titolo esemplificativo, sul piano teorico l'autore dimostra che l'espediente stilistico della modulazione (variazione della chiave) manifesta negli ultimi duecento anni uno spostamento abbastanza costante verso una diminuzione di ridondanza o un aumento di entropia. Laddove Mozart usava soltanto un numero limitato di modulazioni ben definite, Wagner e Debussy ne fecero un uso cosí libero da portare a una perdita quasi completa della significatività della chiave. Molti compositori moderni hanno abolito del tutto il concetto di chiave, tanto che sotto questo profilo la loro musica si avvicina alla completa casualità. Sul piano pratico, Hiller illustra come ha programmato un calcolatore elettronico a comporre musica secondo determinate regole strutturali.

Entropia e ridondanza come misure evolutive. Si possono applicare i risultati della teoria dell'informazione allo studio di organismi viventi. Lila L. Gatlin caratterizza un sistema vivente come sistema di elaborazione d'informazioni, e la vita come «gerarchia strutturale di unità operative, che ha acquisito, attraverso l'evoluzione, la capacità di accumulare ed elaborare le informazioni necessarie per riprodursi in modo fedele» [1972, p. 1]. L'informazione viene immagazzinata nella sequenza delle basi del DNA o, in alcuni casi, dell'RNA.

I sistemi viventi, al pari di altri sistemi che trasmettono informazioni, si trovano di fronte al problema del «rumore». Possono avvenire errori, con eventuale alterazione dell'informazione genetica. Una codificazione sofisticata può servire a ridurre l'effetto del «rumore», ma essa, come si è visto, significa il raggiungimento del punto di equilibrio tra entropia e ridondanza.

Si possono distinguere due metodi per incrementare la ridondanza: il primo consiste nell'uso di segnali non equiprobabili (come già osservato, l'entropia è massima quando tutti i segnali o eventi sono equiprobabili); l'altro consiste nel far dipendere le probabilità dei segnali dalla loro *localizzazione* in una serie di segnali. In un linguaggio naturale si riscontrano entrambi i tipi di vincolo. Nella lingua inglese la lettera *e* è molto piú frequente della lettera *x* (primo metodo) ed è assai probabile che la lettera *q* sia seguita dalla lettera *u* (secondo metodo). Laddove il primo metodo riduce la quantità d'informazione trasmissibile da un dato gruppo di segnali, il secondo contribuisce alla ridondanza e quindi all'affidabilità del codice, senza ridurre la quantità d'informazione potenziale, al solo costo di una maggiore lunghezza dei messaggi.

Nel suo libro la Gatlin dimostra che un'analisi di questo genere può fornire una misura evolutiva. Ella ha calcolato la ridondanza nella sequenza del DNA di vari organismi e ha scoperto che la posizione di un organismo sulla scala evolutiva è funzione diretta della ridondanza introdotta. Secondo quest'analisi, è probabile che l'evoluzione abbia avuto inizio con sequenze casuali di DNA; in seguito, sono stati imposti dei vincoli tali da accrescere l'esattezza dell'informazione genetica. Mentre i sistemi viventi inferiori impiegarono il primo metodo, pagando il prezzo di una ridotta capacità d'immagazzinare informazioni, negli organismi superiori si conseguí una piú elevata ridondanza attraverso il secondo metodo.

3.4. Entropia e informazione in fisica.

Discussa la significatività del concetto di entropia nell'analisi statistica dell'informazione, e stabilita l'analogia formale tra entropia definita secondo la meccanica statistica e secondo la teoria dell'informazione, è ora possibile cercare legami piú sostanziali tra i due concetti di entropia. È abbastanza interessante che la possibilità di una reale connessione tra informazione ed entropia sia stata suggerita molto prima che Shannon proponesse la sua teoria matematica dell'informazione e prima che il concetto di entropia fosse usato al di fuori della fisica.

Il Diavoletto di Maxwell. Nella sua *Theory of Heat* [1871], Maxwell presentò un paradosso che mise a lungo in imbarazzo i fisici. Egli considerò un gas in un contenitore diviso mediante una parete in due scompartimenti A e B. La parete è dotata di un piccolo foro attraverso il quale le molecole del gas possono passare da uno scompartimento all'altro. Il sistema è in equilibrio, ovvero la temperatura è identica nelle due parti. Maxwell ipotizzò inoltre l'esistenza di un Diavoletto capace di sorvegliare i movimenti di ogni singola molecola. Il Diavoletto poteva aprire e chiudere il foro in modo da consentire soltanto il passaggio delle molecole piú veloci da A a B e di quelle piú lente da B a A. Cosí facendo, il Diavoletto, *senza spendere lavoro*, eleverebbe la temperatura di B e abbasserebbe quella di A, in contrasto con la seconda legge della termodinamica.

In un articolo di Szilard [1929] si proponeva di risolvere il paradosso in base alle seguenti considerazioni: il Diavoletto è capace di violare la seconda legge soltanto perché possiede informazioni particolareggiate sullo stato del sistema, informazioni che di solito si suppongono *inottenibili* applicando considerazioni statistiche. Di fatto, il Diavoletto si serve delle informazioni allo scopo di diminuire l'entropia del sistema; in altre parole, secondo Szilard, il Diavoletto *trasforma* informazioni in negentropia. Szilard si serve del termine 'informazione' nel suo senso comune, pre-matematico. La sua soluzione sembra acquistare una maggiore precisione e chiarezza grazie alla definizione shannoniana d'informazione, secondo cui entropia e informazione sono effettivamente espresse dalla stessa funzione matematica. Sembra inoltre che sia sufficiente identificare la costante nella definizione di Shannon con la costante di Boltzmann per consentire di misurare l'informazione in unità termodinamiche. Siffatto pareggiamento di unità è stato criticato per la sua arbitrarietà, soprattutto tenendone presente l'incapacità di funzionare in molti dei casi a cui si può applicare la teoria dell'informazione (e non la termodinamica). Nondimeno, la possibilità stessa del pareggiamento delle unità ha incoraggiato ulteriori sforzi tesi al raggiungimento di un'idea unificata dell'entropia.

Il paradosso di Maxwell è risolto dalle considerazioni di Szilard soltanto a patto che si dimostri che l'azione del Diavoletto produce in realtà piú entropia di quella che va perduta per il processo di classificazione, cosicché *l'entropia dell'intero sistema aumenti*, in accordo con la seconda legge (il Diavoletto dev'essere trattato come essere vincolato alle leggi della fisica; in caso contrario la sua attività si colloca al di là della portata della scienza e non può proporre un problema scientifico). Un'analisi dettagliata dell'azione del Diavoletto è condotta da Brillouin [1962], il quale argomenta che il Diavoletto, per scoprire le singole molecole, deve introdurre nel sistema un agente distruttore dell'equilibrio. Ad esempio, egli può servirsi di una torcia, che rappresenta una fonte di energia «ordinata» o entropia negativa. L'autore dimostra quindi che la seconda legge della termodinamica è sempre soddisfatta a patto di considerare come sistema termodinamico l'insieme costituito dal sistema e dalla fonte esterna di radiazioni.

3.5. Entropia e misura.

Brillouin riesce anche a generalizzare i suoi risultati a tutti i casi di misura fisica. Ogni misura fisica implica un aumento dell'entropia del sistema in cui si realizza la misura. *Ogni guadagno d'informazione* ottenuto con la misura viene quindi *pagato con un aumento di entropia*. Se i nostri risultati sperimentali sono molto precisi, se ne deduce una maggior quantità d'informazione, ma il pedaggio in negentropia è proporzionale all'esattezza dei risultati. Vi sono rigorosi limiti pratici all'accuratezza delle osservazioni, in una certa misura simili alle, ma indipendenti dalle, relazioni d'incertezza nella meccanica quantistica. Una completa accuratezza matematica è in linea di principio irrealizzabile. Proprio questo risultato in particolare è considerato da Brillouin tale da stabilire la connessione tra termodinamica e teoria dell'informazione: informazione ed entropia sono strettamente legate in ogni esperimento realizzato e obbediscono alla seconda legge generalizzata della termodinamica, che pone in relazione guadagno d'informazione e aumento di entropia.

I risultati di Brillouin e di altri operatori del settore indicano che la relazione tra entropia e informazione è piú intima di quanto avrebbe potuto lasciare presupporre il dualismo formale. Pur non potendo pretendere che l'entropia termodinamica *coincida* con l'entropia matematica di un messaggio, si può affermare che – se si riesce a esprimere l'informazione ottenuta da un'operazione fisica (ad esempio una misura) in termini d'incertezza ridotta su una o piú delle variabili fisiche – si stabilisce una relazione matematica invariante tra la quantità d'informazione ricavata e la quantità di entropia prodotta.

È pur vero che nelle discussioni in proposito spesso si sono introdotte molte formulazioni imprecise. Ad esempio si è sovente detto che l'entropia di un sistema fisico è una misura della *nostra mancanza di conoscenza* sul sistema. Ciò equivarrebbe a interpretare l'entropia termodinamica come una quantità del tutto soggettiva, esprimente soltanto una condizione mentale. Vari filosofi, in particolare Karl Popper, hanno a ragione criticato simili interpretazioni soggettiviste. L'entropia di un sistema fisico non è piú soggettiva di qualsiasi altra sua caratteristica fisica. Ciò è confermato dalle definizioni classica e statistica di entropia. Ma Popper sbaglia nel concludere che informazione ed entropia possano essere associate *soltanto* su basi soggettivistiche. La lunghezza di un oggetto fisico è forse la piú oggettiva delle variabili fisiche che possiamo definire. Tuttavia nel misurare la lunghezza, nell'ottenere *informazioni* su di essa, si è vincolati dalla seconda legge generalizzata della termodinamica, formulata da Brillouin, secondo la quale è possibile ottenere informazioni soltanto a spese della negentropia, anch'essa variabile fisica oggettiva. La relazione tra entropia e informazione non è quindi affatto soggettiva, nella misura in cui si evitano formulazioni soggettive inesatte.

3.6. Un'osservazione metodologica.

Un altro rilievo sulla teoria dell'informazione e sulla meccanica statistica fu avanzato da Jayness [1957].

Per raggiungere il suo scopo, cioè la riduzione del comportamento macroscopico a quello microscopico, la meccanica statistica dovette fare varie ipotesi preliminari relative al comportamento dei microsistemi. La teoria dell'informazione ci pone in grado di giustificare alcuni di questi presupposti. Ad esempio: la meccanica statistica dovette attribuire uguali probabilità a priori alle configurazioni fondamentali. A giustificazione dell'ipotesi s'invocò il principio laplaciano della ragione insufficiente, per cui si possono assegnare probabilità uguali ogni qualvolta non vi è motivo per fare diversamente. È chiaro che questo criterio è alquanto insoddisfacente. La riuscita della teoria dell'informazione di Shannon risiede proprio nella sua capacità di darci un criterio oggettivo per la quantità d'incertezza (o entropia) rappresentata da una distribuzione di probabilità. Ne deriva che l'unica attribuzione *razionale* imparziale di probabilità a priori è quella che massimizza l'entropia (l'incertezza) soggette all'informazione di cui si dispone. Laddove nel comune approccio statistico la teoria si andò edificando sulla base di equazioni meccaniche e attribuzione arbitraria di probabilità uguali, e l'identificazione dell'entropia giunse in definitiva per confronto della teoria statistica con la termodinamica classica, prendendo a fondamento la teoria dell'informazione si può iniziare con il concetto di entropia e servirsi della funzione di distribuzione che massimizza l'entropia soggetta a determinati vincoli. La teoria dell'informazione fornisce cosí una *giustificazione metodologica* ai metodi della meccanica statistica.

4. *Aspetti filosofici.*

4.1. La direzione del tempo.

Una delle piú fondamentali intuizioni della mente umana sembra essere il fatto che il tempo si diversifichi dallo spazio sostanzialmente per il possesso di una *direzione*. A differenza dello spazio, considerato isotropo, ossia privo di una direzione preferita, il tempo pare avere un'asimmetria intrinseca. Il percorso da «precedente» a «successivo» è considerato irreversibile. Vi sono due diverse interpretazioni filosofiche di questa fondamentale intuizione: la si può ritenere un aspetto soggettivo del modo umano di concepire il fenomeno, privo di realtà oggettiva, oppure la si può considerare una caratteristica intrinseca della natura. Benché il secondo punto di vista sia piú diffuso tra i profani come tra gli scienziati, alcuni eminenti filosofi aderirono al primo. Spinoza, ad esempio, sostenne che «tempus non est affectio rerum sed merus modus cogitandi». Se si accetta la seconda prospettiva, vale a dire se si dà alla differenza tra tempo e spazio il valore di fenomeno oggettivo, si possono

ancora adottare due atteggiamenti diversi: valutare l'idea di tempo come concetto primario, del quale l'asimmetria o la direzione è una caratteristica, o cercare di *ridurre* la direzione del tempo ad altri concetti piú semplici o meglio conosciuti. Per effettuare questa riduzione è necessario scovare un principio implicante una fondamentale asimmetria o anisotropia, ed è a questo punto che entra in ballo il concetto di entropia. Come si è notato, l'idea che i fenomeni naturali puntino in una certa direzione si materializza nella seconda legge della termodinamica. Irreversibilità e aumento di entropia sono concetti inseparabili nella spiegazione classica come in quella statistica. Sembra perciò del tutto naturale che i tentativi di ridurre la direzione del tempo ad altri principî fisici si siano appellati alla legge dell'aumento dell'entropia.

L'identificazione della direzione del tempo con la direzione della variazione di entropia ebbe origine negli scritti di Boltzmann e in seguito è stata raccolta e ridotta a sistema da vari filosofi, tra cui Arthur Eddington nel suo libro *The Nature of the Physical World* [1928]. Eddington propone di definire la relazione «piú tardi di» in termini di aumento di entropia: di due stati, quello con entropia maggiore è il successivo. La «freccia del tempo» (metafora di Eddington per descrivere l'asimmetria tra direzioni temporali positiva e negativa) si spiega, o si riduce, alla «freccia» dell'entropia crescente.

Si presentano immediatamente varie difficoltà:

1) L'entropia è sempre definita per un particolare sistema fisico. La direzione del tempo, secondo la definizione di Eddington, si applicherà perciò anche a sistemi isolati, perdendo cosí apparentemente la sua generalità. È possibile sormontare tale difficoltà definendo la direzione del tempo sulla base dell'aumento di entropia nell'intero universo, ma non è facile giustificare una simile estrapolazione del concetto di entropia (cfr. oltre, § 4.3). Inoltre, si supponga di avere due sistemi A e B e si ipotizzi che lo stato entropico $S(A)_1$ sia contemporaneo allo stato $S(B)_1$. Se si verifica il caso di uno stato $S(A)_2$ contemporaneo a uno stato $S(B)_2$ e tale che $S(A)_2 > S(A)_1$, ma $S(B)_2 < S(B)_1$, allora la definizione della direzione del tempo è inconsistente. È quindi necessaria una prova della consistenza della nostra definizione.

2) La seconda legge della termodinamica sembra perdere la sua connotazione di legge fisica, riducendosi a una tautologia. Dato il concetto di tempo anisotropo, la seconda legge ha un contenuto chiaro: essa afferma che l'entropia aumenta col tempo, ovvero predice che se fosse possibile osservare un sistema dopo un certo tempo, probabilmente lo si troverebbe in uno stato di entropia maggiore. Controprove alla legge saranno sistemi caratterizzati da un'entropia in costante *diminuzione* col tempo. Ma, se si definisce la relazione «piú tardi di» in termini di aumento di entropia, non vi sono leggi, né previsioni, né controprove possibili. Gli stati di entropia piú bassa dovranno essere considerati precedenti *per definizione*.

È possibile rispondere parzialmente a questa obiezione in modo da risolvere anche il problema della consistenza. Si può attribuire alla seconda legge della termodinamica il valore di principio empirico che garantisce la consistenza della definizione della direzione del tempo, Cosí la definizione è salva e la legge ha ancora un contenuto empirico. È opinabile se questa costruzione filosofica migliori realmente la nostra comprensione della freccia del tempo.

3) Non tutti i processi coinvolgono un aumento di entropia. I processi reversibili non alterano l'entropia del sistema in cui hanno luogo. Un sistema già in equilibrio non aumenterà ulteriormente la sua entropia. La definizione di Eddington implica che per questi processi il tempo sia isotropo o – per usare le sue parole – perda la sua freccia. Ancora una volta la conclusione è contraria all'intuizione e getta dei dubbi sul valore della definizione.

4) Eddington pensava che l'aumento di entropia fosse l'unico fenomeno naturale in grado di assicurare un terreno solido per la determinazione della direzione del tempo. Per tutte le branche della fisica, termodinamica esclusa, un asse isotropo del tempo è sufficiente. Tuttavia alcuni hanno rilevato (ad esempio Lewis, Popper, Costa de Beauregard) che l'irreversibilità non è caratteristica esclusiva della termodinamica. Un tipico esempio di irreversibilità al di fuori della termodinamica è la propagazione di un'onda in circoli concentrici attorno a una pietra gettata in uno stagno. Il processo inverso – un'onda di contrazione – non si osserva in quanto le condizioni che lo provocherebbero sono molto improbabili. Sulla scorta di esempi di questo tipo, taluni autori traggono la conclusione che l'irreversibilità non è un elemento sufficiente per la definizione della direzione del tempo, mentre altri sostengono che la seconda legge della termodinamica è soltanto una delle varie manifestazioni dell'anisotropia del tempo. In entrambi i casi si rifiuta la relazione esclusiva tra entropia e freccia del tempo.

5) La quinta difficoltà è forse la piú grave. Si è visto che la seconda legge della termodinamica viene oggi considerata una legge statistica. L'entropia di un sistema chiuso subisce fluttuazioni: è ben lungi dal crescere sempre. L'anisotropia del tempo, d'altra parte, almeno secondo la nostra concezione intuitiva, non sembra avere fondamenta statistiche. La riduzione appare quindi insoddisfacente. Per apprezzare la forza di questo argomento si riesamini il paradosso di Zermelo, discusso nel § 2: la conclusione è stata che il paradosso non invalida la seconda legge e che non si ha contraddizione tra l'opinione di Boltzmann, secondo la quale gli stati di entropia elevata sono piú probabili, e il teorema della periodicità di Poincaré. Tuttavia il paradosso costituisce sempre un ostacolo alla definizione entropica del tempo. Un sistema chiuso deve alla fine ritornare a un determinato stato iniziale, e quindi non si evolve continuamente nella medesima direzione. Zermelo argomenta cosí che nella misurazione dell'entropia di un sistema è impossibile dedurre che il sistema si sia in precedenza trovato in uno stato di entropia piú bassa. È possibile che il sistema avesse prima un'entropia maggiore e che ora si trovi in fase di diminuzione entropica, ritornando a un certo stato iniziale. Ne consegue l'impossibilità di dedurre la direzione del tempo dal comportamento entropico del sistema.

Reichenbach [1956] e Grünbaum [1963], difensori della definizione entropica di anisotropia, si dànno entrambi un gran da fare per rispondere a quest'argomentazione. Trascurando molti dei particolari della discussione, l'idea centrale della loro soluzione è la seguente: essi ammettono che la dire-

zione del tempo *non* sia deducibile dal comportamento di un sistema *chiuso*, ma al contempo sostengono che i sistemi fisici reali sono soltanto quasi-chiusi e solo per periodi di tempo relativamente brevi; prima o poi essi interagiscono con l'ambiente esterno. Osservando un sistema fisico reale possiamo perciò ignorare la possibilità che abbia rivestito il carattere di sistema chiuso per un periodo sufficiente a completare un ciclo secondo quanto richiede il teorema di Poincaré. Se ad esempio accade d'imbattersi nottetempo in una roccia calda, è molto piú ragionevole dedurre che essa sia stata in precedenza riscaldata grazie all'*interazione con un altro sistema* – il Sole – e che attualmente sia in fase di *raffreddamento*, piuttosto che supporre di osservare una fluttuazione in un sistema chiuso e un processo di diminuzione dell'entropia della roccia. Per quanto l'argomento sia valido, e benché Eddington avesse ragione a far rilevare che in presenza dei grandi numeri il caso è la migliore garanzia di certezza, è nondimeno possibile porre in dubbio, insieme a Carnap, la legittimità dell'assunzione di una corrispondenza statistica, per quanto elevata, come base per una definizione teorica.

In considerazione di tutte le difficoltà con cui si scontra la definizione entropica della direzione del tempo, sembra ragionevole trarre la seguente conclusione: sinora non è stata conseguita una riduzione soddisfacente della freccia del tempo alla freccia dell'entropia. È comunque possibile difendere una tesi filosofica piú limitata, nel senso che l'aumento di entropia costituisce la migliore manifestazione della freccia del tempo nel mondo fisico. In termini piú pittoreschi: un uomo che perdesse improvvisamente la coordinazione nel tempo, il senso del passato e del futuro, potrebbe adoperare stati di entropia crescente come segnali indicatori per riorientarsi, e cosí facendo con ogni probabilità si troverebbe in ottimo accordo col resto dei membri della sua società, che si servono della propria freccia interna del tempo.

4.2. Entropia ed evoluzione.

I sistemi biologici sembrano a prima vista violare la seconda legge della termodinamica. Essi sono capaci di conservare le proprie strutture estremamente ordinate e di riprodurle con grande precisione; non si evolvono nella direzione del disordine crescente o entropia crescente. Tuttavia questa contraddizione è soltanto apparente. I sistemi biologici sono sistemi *aperti* che scambiano costantemente energia con l'ambiente. Si deve calcolare il bilancio totale dell'entropia nel sistema chiuso, includendo sia l'organismo sia l'ambiente; ma in questo caso si può dimostrare che la seconda legge viene rigorosamente osservata. Una diminuzione locale di entropia in un punto del sistema è compatibile con la seconda legge nella misura in cui è equilibrata da un aumento uguale o maggiore in altre parti del sistema. La cristallizzazione di una soluzione satura costituisce un esempio di diminuzione locale di entropia. La perdita di entropia nella cristallizzazione è bilanciata da un passaggio di calore dalla fase cristallina alla soluzione. L'entropia del sistema, come entità globale, aumenta.

Si ritiene che lo stesso criterio valga per i sistemi viventi. Di fatto si può calcolare il bilancio entropico in alcuni casi semplici, per i quali si può *dimostrare* l'assenza di violazione della seconda legge. Ad esempio è possibile allevare batteri in una soluzione di glucosio e prodotti minerali. Dopo alcune ore la soluzione conterrà un gran numero di batteri. Durante il processo, una parte dello zucchero si trasforma in costituenti cellulari (ossia diminuisce la propria entropia), mentre il resto si ossida ad anidride carbonica e acqua (ovvero aumenta la propria entropia). Facendo un dettagliato bilancio entropico (a questo scopo si deve effettuare l'intero esperimento in un calorimetro), si può dimostrare che l'entropia del sistema globale in effetti subisce un aumento. La riproduzione di un sistema vivente non offre perciò prove contrarie alla seconda legge.

Un problema piú intricato è sollevato dal fenomeno dell'evoluzione in generale. L'evoluzione organica sembra costituire un processo irreversibile indirizzato verso una maggiore complessità e un piú elevato ordine. Non è forse una prova sensazionale contro la validità generale della seconda legge? Non esiste una risposta empirica immediata al problema. Non si dispone di mezzi per valutare il bilancio entropico dell'intera storia dell'evoluzione. È però possibile aggirare la difficoltà: il punto di partenza dei cambiamenti evolutivi biologici è rappresentato dalle *mutazioni* nella struttura del DNA (si tralascia in questa sede il problema della creazione del primo sistema vivente). In molti casi queste mutazioni si sono dimostrate *reversibili*. Il meccanismo che introduce l'irreversibilità nel processo dell'evoluzione, secondo la teoria di Darwin, è la selezione naturale. Non possiamo quindi vedere nell'ordine crescente lungo il percorso evolutivo l'espressione di una tendenza ad aumentare l'ordine. Soltanto attraverso la *selezione* alcune mutazioni persistono e altre scompaiono, senza che venga in alcun modo violata la seconda legge. Questa risposta certamente soddisfa il riduzionista che crede nella riduzione delle leggi biologiche a leggi fisiche. Tuttavia alcuni biologi e filosofi rifiutano di accettare l'ipotesi e credono nell'esistenza di «leggi biotoniche» che, tra le altre cose, governano l'evoluzione mediante meccanismi diversi dalla selezione naturale. Esula da questo articolo la discussione approfondita di queste opinioni.

4.3. Morte termica dell'universo.

Si è ripetutamente fatto rilevare che il principio di aumento dell'entropia si applica esclusivamente ai sistemi chiusi (quello di sistema chiuso è naturalmente uno dei concetti idealizzati della scienza, nella realtà pratica solo approssimabili). Il problema che ha reso perplessi gli scienziati sin da quando Clausius annunziò il suo principio è il seguente: si può considerare l'intero universo come sistema chiuso? Si può supporre che la seconda legge della termodinamica significhi che l'entropia dell'universo tende a un massimo? Rispondendo positivamente, come fece lo stesso Clausius, ne consegue che l'energia da un capo all'altro dell'universo subisce una degradazione e che avremo a disposizione una quantità via via minore di energia utile. Lo stato finale

di entropia massima provocherebbe quello che Clausius definí *Wärmetode*, la «morte termica» dell'universo. L'unica forma di energia esistente in quello stato sarà energia termica a temperatura costante, inutile quanto una completa assenza di energia. In queste condizioni la vita sarebbe assolutamente impossibile.

Alcuni fisici contemporanei di Clausius rifiutarono di accettare una visione cosí severa, basata su un'estrapolazione del concetto di sistema chiuso. In una breve lettera al «Philosophical Magazine», nel 1852, William Rankine propose che la degradazione dell'energia nel nostro settore di universo si accompagni a un processo di riconcentrazione o rifocalizzazione dell'energia in qualche altra parte.

Nonostante i grandi sviluppi in astronomia e in astrofisica, la questione non è ancora risolta. Le prove paiono aperte a piú di un'interpretazione. Ad esempio, il modello dell'universo in espansione, apparentemente confermato dallo spostamento verso il rosso della luce che perviene alla Terra da stelle distanti, da un lato è coerente con un inizio definito dell'universo nel tempo a partire da uno stato di bassa entropia, ma dall'altro lascia spazio ad argomentazioni sulla scarsa legittimità del considerare l'universo come un sistema chiuso [Katchalsky e Curran 1965].

Il problema non è affatto semplificato dall'interpretazione statistica dell'entropia. Si può ragionare sull'eventuale salvezza dell'universo dalla morte termica grazie a una fluttuazione dallo stato di massima entropia, ma è impossibile esprimere la probabilità di una fluttuazione abbastanza lunga da continuare o sviluppare nuovamente la vita. È possibile che il nostro intorno di universo sia un'ampia fluttuazione in un universo in equilibrio e che – ristabilito l'equilibrio nella nostra zona – abbia inizio altrove un'altra fluttuazione. Al momento ci si deve limitare alle speculazioni: una teoria termodinamica scientifica dell'intero universo sembra fuori portata.

4.4. Il concetto di entropia e la valutazione della scienza.

A conclusione dell'articolo, è opportuna una breve rassegna dell'evoluzione del concetto di entropia come capitolo nella storia della scienza. Come si è visto, il concetto emerse dagli studi di un ingegnere interessato ai motori termici. All'epoca di Carnot non si poteva immaginare che le sue *Réflexions*, per quanto brillanti, avrebbero dato vita a un principio che si colloca tra quelli scientifici fondamentali. Tuttavia non ci volle molto per rendersene conto. Con Clausius, i risultati di Carnot sono già elevati al rango di principio allineato al principio di conservazione dell'energia. Il principio è divenuto il fondamento di una nuova branca della scienza, la termodinamica. Tuttavia la termodinamica, al pari di qualsiasi altro ramo della scienza, non poteva restare una teoria isolata, ma doveva essere collegata ad altre teorie fisiche. Essa richiedeva fondamenti piú particolareggiati e doveva essere applicata a nuovi problemi. Cosí, in conseguenza della ricerca di un legame tra termodinamica e meccanica, si sviluppò la teoria statistica dell'entropia. Il concetto di entropia subí un cambiamento radicale nel corso del suo sviluppo. L'entropia è divenuta una misura statistica del grado di ordine di un sistema e il principio di aumento dell'entropia ha acquisito le caratteristiche di un principio statistico. Le nuove basi della teoria crearono nuove applicazioni e nuovi problemi e la termodinamica, nella sua nuova forma, divenne oggetto delle attenzioni piú del teorico che dell'ingegnere. È abbastanza interessante il fatto che un nuovo capitolo nella storia del concetto abbia avuto origine con un semplice saggio intitolato *The Mathematical Theory of Communication* [Shannon e Weaver 1949] ad opera di uno scienziato rivolto soprattutto a problemi tecnici. Il concetto di entropia si è nuovamente trasformato in un altro concetto, identico nel nome, ma diverso nel contenuto che esprime. Di nuovo si dovettero stabilire legami tra il nuovo concetto di entropia e quelli vecchi. Alcuni di questi legami sono soltanto formali, altri piú sostanziali. E, come già accaduto in precedenza, il nuovo contenuto del concetto ha aperto un intero campo di nuove applicazioni [cfr. Kubát e Zeman 1975].

Il concetto di entropia non è l'unico concetto scientifico che nell'evoluzione storica abbia subito cambiamenti e trasformazioni. La storia della scienza abbonda di altri esempi. Due caratteristiche sembrano tipiche dell'evoluzione dei concetti scientifici fondamentali: da un lato vi è il processo dell'unificazione di gruppi sempre piú ampi di fenomeni nello stesso schema concettuale; dall'altro diviene chiaro che ogni *estensione* di una teoria implica un *mutamento* della teoria stessa, e che occupandosi della storia di un concetto scientifico si ha in realtà a che fare con un'intera famiglia di concetti che, come i membri di una stessa famiglia, portano lo stesso cognome, hanno una certa somiglianza, ma nondimeno non sono copie dello stesso individuo. Ciò che Wittgenstein ha sottolineato in relazione ai concetti del nostro linguaggio ordinario vale certamente anche per i concetti scientifici. [Y. E. e Y. B.-M.].

Boltzmann, L.
1896-98 *Vorlesungen über Gastheorie*, Barth, Leipzig.

Brillouin, L.
1962 *Science and Information Theory*, Academic Press, New York 1962².

Carnot, S.-N.-L.
1824 *Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance*, Bachelier, Paris.

Clausius, R.
1864-67 *Abhandlungen über die mechanische Wärmetheorie*, 2 voll., Vieweg, Braunschweig.

Eddington, A. S.
1928 *The Nature of the Physical World*, Cambridge University Press, Cambridge.

Fast, J. D.
1959 *Entropie; de betekenis van het entropiebegrip en zijn toepassingen in wetenschap en technick*, Centen, Hilversum 1959²; trad. ted. Philips' Technische Bibliothek, Eindhoven 1960.

Gatlin, L. L.
1972 (a cura di) *Information Theory and the Living System*, Columbia University Press, New York and London.

Gibbs, J. W.
1902 *Elementary Principles in Statistical Mechanics*, Scribner, New York; ora in *Collected Works*, vol. II, parte I, Yale University Press, New Haven Conn. 1948.

Grünbaum, A.
1963 *Philosophical Problems of Space and Time*, Knopf, New York.

Hiller, L. A.
1959 *Computer Music*, in «Scientific American», CCI, n. 6, pp. 109-21.

Jayness, E. T.
1957 *Information Theory and Statistical Mechanics*, in «Physical Review», CVI, pp. 620-30.

Katchalski, A., e Curran, P. F.
1965 (a cura di) *Nonequilibrium Thermodynamics in Biophysics*, Harvard University Press, Cambridge Mass.

Kubát, L., e Zeman, J.
1975 (a cura di) *Entropy and Information in Science and Philosophy*, Elsevier, Amsterdam.

Maxwell, J. C.
1871 *Theory of Heat*, Longmans and Green, London.

Reichenbach, H.
1956 *The Direction of Time*, University of California Press, Berkeley and Los Angeles.

Shannon, C. E., e Weaver, W.
1949 *The Mathematical Theory of Communication*, University of Illinois Press, Urbana Ill. (trad. it. Etas Kompass, Milano 1971).

Szilard, L.
1929 *Über die Enthropieverminderung in einem thermodynamischen System beim Eingriffen intelligenter Wesen*, in «Zeitschrift für Physik», LIII, pp. 840-56.

Mentre la meccanica classica (cfr. **fisica**), fondata su leggi di **conservazione/invarianza**, supponeva un'**equivalenza** nell'ordine di produzione dei fenomeni (cfr. **fenomeno**), in tutti i processi unidirezionali (cfr. **reversibilità/irreversibilità**), come certi fenomeni di **propagazione**, l'entropia non può che aumentare (cfr. **ordine/disordine**). Ciò significa che gli stati di entropia piú elevata sono piú probabili (cfr. **probabilità**) di quelli con entropia meno elevata, ovvero che c'è equilibrio (cfr. **equilibrio/squilibrio**) quando un **sistema** oscilla intorno allo stato piú probabile (cfr. **stato fisico**). L'entropia, interpretata dapprima in termini finalisti, ha potuto essere spiegata (cfr. **spiegazione**), grazie alla meccanica statistica, a partire dal **moto** delle particelle (cfr. **particella, atomo e molecola**). In seguito, assegnando una funzione d'entropia a ciascuna distribuzione di probabilità (cfr. **distribuzione statistica**) indipendentemente dai contesti fisici, si è costruita una teoria (cfr. **teoria/modello**) dell'**informazione** in cui l'entropia ne è la **misura** – l'informazione cresce con l'entropia e, inversamente, ogni **vincolo** comporta una perdita d'informazione (cfr. **comunicazione, codice, linguaggio**) – ed è stato possibile interpretare i sistemi viventi (cfr. **vita**) come un sistema di trattamento dell'informazione (cfr. **eredità, gene**). L'importanza esplicativa del **concetto** di entropia è in continua crescita (cfr. **paradigma**): si pensi al ruolo considerevole che ha assunto nelle attuali discussioni sulla direzione del tempo (cfr. **tempo/temporalità, ciclo**), sull'**organizzazione** e sull'**evoluzione** del vivente e sulla «morte termica» dell'**universo** (cfr. **cosmologie, energia**).

Fisica

1. *Generalità.*

Secondo le vecchie definizioni le scienze della materia inanimata erano essenzialmente due, la fisica e la chimica. La prima trattava dei fenomeni naturali su grande scala, la seconda riguardava le relazioni tra corpi puri (o «sostanze chimiche»), in particolare le loro trasformazioni gli uni negli altri.

L'evoluzione delle conoscenze ha reso antiquata questa suddivisione. Oggi, in conformità coll'etimologia, per fisica (dal greco φύσις 'natura') s'intende in generale la scienza che ha per oggetto l'insieme di quei fenomeni che comunemente vengono detti naturali, ivi compresi quindi anche quelli chimici. La nozione generale di fisica comprende dunque attualmente, oltre alla chimica, anche la meccanica – a torto considerata in precedenza come un ramo della matematica –, nonché la maggior parte dell'astronomia (astrofisica). Ne restano fuori, con lo status di scienze distinte e autonome, alcune discipline di carattere essenzialmente descrittivo come la geologia e le scienze della vita e della coscienza. Per quel che riguarda queste ultime, non ci si può ancora pronunziare nemmeno sulla questione se i loro principî siano o meno riducibili a quelli della fisica intesa nel senso appena specificato. La questione è sicuramente una delle piú appassionanti, ma le indicazioni di cui si dispone sono ancora troppo frammentarie – e insieme troppo ambigue – per autorizzare una qualsiasi presa di posizione su questo punto; in realtà esse non autorizzano nemmeno ad affermare che la questione sia ben posta, come si potrebbe constatare proprio in questo articolo.

La definizione di fisica appena data si riferisce – come si può vedere – all'oggetto studiato; ma, come si può facilmente constatare, coincide all'incirca con quella che se ne potrebbe dare in relazione al metodo. Tra le scienze empiriche la fisica è quella che piú delle altre utilizza la matematica – e deve utilizzarla – nella formulazione delle proprie leggi.

Le definizioni e le asserzioni che precedono suscitano indubbiamente un certo numero di domande. La prima, forse la piú importante, quella piú di fondo comunque sotto il profilo filosofico, si può formulare cosí: «Cos'è infine la Natura?» Ovvero: «Che cos'è un *fenomeno* naturale?» È scontato che l'Uomo non goda di una conoscenza *diretta* della realtà. Per cercare di conoscerla deve dunque *osservare* e, se mira alla precisione, deve ***misurare***. Cosí, quel che conosce con certezza, come i filosofi a buon diritto hanno da un pezzo sottolineato, è esclusivamente il risultato di tali osservazioni o di tali misure. Come andrà dunque concepita una scienza della Natura? Una domanda del genere pone, evidentemente, il problema della relazione tra l'uomo e la realtà. Si tratta di un problema antico, che agli albori delle moderne scienze matematiche Descartes aveva colto con chiarezza, ma che i fisici riuscirono però per lungo tempo ad evitare: cosa che oggi, con il progresso delle

conoscenze riguardanti la struttura intima della materia, non sembrano piú poter fare.

Una seconda domanda che sorge con la nostra definizione della fisica è: «Pur ammettendo che ci sia una concezione accettabile della nozione di Natura e che si sia trovato un accordo sulla nozione di "scienza della Natura" in generale, in che misura sussiste davvero l'*unità* di questo sapere circa la natura»?

Una terza domanda riguarda infine il ruolo della matematica nella fisica. Si è visto come un tal ruolo sia essenziale e, per cosí dire, «costitutivo»: ma in che cosa consiste esattamente?

Ogni analisi della fisica contemporanea deve tenere presente sin dall'inizio l'esistenza di queste tre domande: bisogna essere consapevoli del fatto che tutte e tre sono domande complesse e ammettono solo risposte sfumate: altrimenti c'è il rischio di partire sulla base d'idee false e di dover poi procedere a difficili revisioni concettuali. D'altra parte le risposte a queste domande non vanno cercate a priori. Conviene procedere guardando alla fisica quale è oggi. In questo articolo si affronteranno quindi tali domande solo in un secondo tempo (la seconda e la terza nel § 3; la prima, che è la piú difficile, nel § 4).

2. *Breve profilo storico della fisica e del suo ruolo nell'evoluzione delle idee.*

In un articolo come questo non si può presumere di riuscire a puntualizzare in modo imparziale e circostanziato, coi metodi dello storico delle idee, lo sviluppo nel corso dei secoli della fisica nel suo complesso. L'argomento è troppo ampio. D'altra parte, non c'è dubbio che un'autentica comprensione della natura dei problemi che riguardano la relazione Uomo-Natura, cosí cruciale per la fisica contemporanea, possa essere in larga misura facilitata dalla conoscenza preliminare del modo in cui questi problemi sono lentamente maturati nella mente dei pensatori, filosofi o scienziati che fossero, óppure l'una cosa e l'altra insieme (caso che si verificò di frequente). O meglio, pare evidente che a tale comprensione gioverà uno studio, sia pur succinto, dei modi in cui le idee influenzarono nel corso dei secoli le ricerche e i risultati ottenuti con tali ricerche modificarono a loro volta tali idee. Questo paragrafo vuole appunto delineare un abbozzo, non esauriente, della storia di tali rapporti.

2.1. L'antichità.

È risaputo che le civiltà orientali, in primo luogo quella cinese, hanno fornito importanti contributi alla fondazione della scienza e della tecnica. Anche la civiltà maya aveva da parte sua ottenuto, nel campo della misurazione del tempo e dell'astronomia di posizione, risultati assai significativi. E tuttavia nel complesso pare giusto, come fa la tradizione, far risalire al pensatore ionico Talete l'inizio tanto della scienza quanto della filosofia (è noto che per tutta l'antichità scienza e filosofia non furono affatto divise).

Si usa dire che è la scoperta, certamente ancora confusa, dell'idea di *legge naturale* a distinguere dai predecessori Talete e i pensatori della sua scuola. Prima di loro le spiegazioni del mondo si configuravano, in modo piú o meno netto, come *genesi*, cioè come racconti fondati su nozioni radicalmente antropomorfe. In questi racconti la natura delle cose e della loro totalità, l'universo, era spiegata con un semplice riferimento all'esistenza di una o di molteplici volontà sovrumane. Nella civiltà occidentale un'idea diversa e alternativa rispetto a questa su basi antropomorfiche appare per la prima volta, sembra, nella Ionia ed è appunto quella che consiste nel conferire un'importanza particolare alla nozione di una sostanza universale e permanente, non specificamente «spirituale», ma dotata di strutture anch'esse permanenti o per lo meno stabili. Mancavano allora, beninteso, i termini atti a esprimere concezioni di un tal livello di astrazione, che di conseguenza non furono in grado di affermarsi immediatamente con chiarezza. Ancora prigioniero di una concretezza che ci appare oggi ingenua, Talete affermava che «tutto è acqua». Ma assai presto dopo di lui comparve la dottrina dei quattro elementi e della tendenza spontanea dei corpi a tornare al luogo naturale. Quest'idea, cara ad Aristotele, doveva poi svolgere a lungo un ruolo di notevole importanza. Se nei dettagli riveste per noi scarso interesse, dal punto di vista storico ha però il merito sostanziale di costituire uno dei primi esempi conosciuti di teoria scientifica. Si tratta cioè di un'idea generale da cui è possibile trarre deduzioni che si possono applicare ai fatti osservati e che comportano implicazioni verificabili: se queste ultime si rivelano contrarie all'esperienza, la teoria può essere allora scartata. Fu quel che effettivamente accadde con la teoria dei quattro elementi, anche se solo molto piú tardi.

Il sorprendente pensiero della scuola ionica aveva intanto continuato il suo sviluppo. La sua storia – e va notato, perché anche questo ha la sua importanza – mette in luce l'angustia intellettuale di quei nostri esteti che, abbagliati dalla potenza industriale dei tempi moderni, non sognano altro che cemento armato e ingranaggi, ostentando un divertito disprezzo per la grazia e la bellezza. Coloro che per questo motivo si mostrano tanto sprezzanti verso la «grazia ionica», in realtà ignorano che a questa Ionia «tanto aggraziata» e ai suoi dintorni la nostra civiltà è debitrice proprio della sua potenza tecnologica! Quest'ultima si fonda infatti sia sulla matematica astratta – creata, come sembra, da Talete e da Pitagora di Samo – sia sulla fisica, che affonda le sue radici teoriche nel pensiero di Leucippo e Democrito, come pure in quello stesso di Pitagora e dei suoi discepoli della Magna Grecia. Si tornerà in seguito sulle idee-forza dell'atomismo (cioè sulla dottrina di Leucippo e Democrito) e su quelle del «matematismo», che deriva dall'atteggiamento intellettuale di Pitagora e anche, seppur in modo piú sfumato, da quello di Platone. Si tratta di due concezioni estremamente differenti e, sotto un certo profilo, addirittura opposte, alle quali la fisica deve pressoché tutto, anche se oggi deve in qualche misura oltrepassare sia l'una che l'altra. Ci si può qui limitare alla congettura che la loro origine nel paese della grazia e dell'equilibrio non sia un puro capriccio del caso, ma riveli forse una profonda unità dello spirito umano nascosta sotto l'estrema diversità dei suoi percorsi piú fecondi.

Salvo rare eccezioni, i Greci non praticarono la sperimentazione. L'idea di creare artificialmente certe condizioni per isolare meglio un certo fenomeno sembrava loro quasi una forma di violenza sulla Natura, tanto da considerarla incompatibile con una sana concezione della misura che l'uomo deve tenere nelle proprie azioni. L'assenza pressoché totale di sperimentazione colpisce oggi come la lacuna piú sorprendente della scienza greca. I Greci facevano però delle osservazioni. Aristotele in particolare è da questo punto di vista un notevole precursore, benché, soprattutto, per quel che concerne la biologia. Ma anche nel campo dell'astronomia l'osservazione portò a misure abbastanze precise. A loro volta queste, congiunte a principî generali accettati per l'evidente semplicità e per l'appagamento che la loro eleganza procurava allo spirito, dettero origine a cosmologie. Tra queste ultime va segnalata quella di Aristarco di Samo, che era eliocentrica: secondo questo precursore di Copernico, è la Terra che gira intorno al Sole e non il contrario. L'ipotesi non fu tuttavia recepita dai suoi contemporanei e, com'è noto, fu proprio la teoria opposta, quella geocentrica, opportunamente rielaborata da Tolomeo d'Alessandria in modo da render conto delle misure ottenute per i pianeti, che in definitiva s'impose per molti secoli.

Nell'antichità la scienza non risulta nettamente distinta dalla filosofia. In altri termini, le idee scientifiche sono fortemente condizionate dai presupposti filosofici. A questo riguardo, Platone è stato sovente fatto oggetto di vivaci critiche. Son critiche meritate nella misura in cui, convinto come era che da una parte le Idee fossero per definizione le sole entità autenticamente reali e dall'altra che esse fossero direttamente accessibili al pensiero, Platone giungeva a considerare come fallace la realtà sensibile e di conseguenza la sua osservazione come fonte d'inganno, quindi inutile, anzi dannosa. Ma da un altro punto di vista il giudizio circa l'influenza che il pensiero platonico esercita sulla fisica non può limitarsi alla sola critica in quanto proprio tale pensiero ha reso possibile un distacco dai dati immediati che sul lungo periodo si è rivelato spesso benefico. Senza questo distacco, lo spirito umano avrebbe potuto essere indebitamente indotto a tramutare l'oggetto sensibile in un assoluto e quindi a esercitare la propria forza immaginativa nei ristretti limiti di quanto può essere concepito con chiarezza sulla base dell'esperienza quotidiana. Nelle sue memorie Heisenberg, giudicando il fatto non privo di significato, ricorda che quando, ancora studente, cominciò ad interessarsi al problema della struttura degli atomi e della natura dei loro legami, i *Dialoghi* di Platone erano tra le sue letture preferite. Proprio il pensiero platonico l'ammoniva a non cadere nella tentazione d'immaginare come reali i ganci che i manuali di chimica descrivevano con indulgente cura e che rappresentavano i legami atomici. Senza dubbio i *Dialoghi* gli suggerivano anche che una rappresentazione matematica piú astratta di questi legami aveva maggiori probabilità di corrispondere alla realtà profonda.

Un riepilogo del pensiero greco nelle sue relazioni con la fisica contemporanea, per quanto sintetico, non può infine non ricordare un outsider: il sofista Protagora. Una frase enigmatica che egli ha scritto, «L'uomo è misura

di tutte le cose», sembra farne un precursore dei positivisti, il cui ruolo nello sviluppo della fisica contemporanea è stato, come si vedrà, essenziale.

2.2. Rinascimento ed età classica (secoli XVI, XVII, XVIII).

Nel medioevo, epoca assai feconda di progressi tecnici, i pensatori si sforzarono di approfondire entro una visione cristiana del mondo i grandi problemi filosofici ereditati dai Greci. Questi sforzi speculativi dettero luogo a controversie rimaste celebri; ma alla fine di quest'epoca era ormai divenuto chiaro a molti che nessun progresso veramente decisivo ne era in definitiva risultato. Va attribuito almeno in parte al senso di stanchezza che si accompagnò a questa situazione di fatto il rinnovarsi dell'interesse, manifestatosi nel Rinascimento, per le questioni legate all'esistenza terrena e quindi per la Natura nel cui seno tale esistenza si sviluppa. È a partire dal XVI secolo che non pochi pensatori di primo piano cominciano a volgersi al particolare e al concreto: il lavoro dell'artigiano non è piú sottovalutato come nelle epoche precedenti. In concomitanza a questa nuova situazione si riprende l'osservazione dei fenomeni naturali. E si comincia soprattutto a fare veri esperimenti.

Per prima cosa, l'osservazione. Questa si esercita principalmente sulle stelle e sui pianeti, campo nel quale è fortemente stimolata dalle credenze astrologiche che fioriscono in questo periodo. È soprattutto Tycho Brahe che, mettendo a profitto i piú recenti progressi tecnici, riesce a ottenere nel suo osservatorio misure di alta precisione. E sono queste misure che permettono a Keplero di perfezionare l'ipotesi eliocentrica poco tempo prima riscoperta da Copernico. Le orbite della Terra e dei pianeti intorno al Sole si rivelano ellissi e non cerchi. Questa perdita di semplicità, dunque di eleganza apparente (una ellisse è «meno bella» di un cerchio, avrebbero detto i Greci), è tuttavia compensata dal fatto che le caratteristiche delle traiettorie in questione, cosí come quelle dei corpi celesti che le percorrono, obbediscono a regole semplici, quantitative, esprimibili matematicamente. Da allora questo diventa il vero criterio per giudicare la semplicità e l'eleganza di una teoria, e pertanto si afferma progressivamente come l'indicazione a cui il fisico tende a conferire il peso maggiore nella ricerca di una teoria «vera».

Passando ora alla sperimentazione, è ovvio che si comincia citando Galileo. Egli non fu però un isolato. Come lui, o seguendo il suo esempio, parecchi ricercatori non esitano in questo periodo a specializzarsi nello studio quantitativo, addirittura meticoloso, di uno o piú problemi specifici, malgrado tali problemi rivestano ancora scarsa importanza agli occhi dei filosofi. È in questo periodo che nasce veramente la fisica. Ben presto questi «specialisti» scoprono con loro grande meraviglia non solo che il loro modo di procedere li mette in grado di trovare risposte sicure, definitive, prive di contraddizioni ai limitati problemi che si sono posti, ma che anche il campo di tali problemi si allarga progressivamente e che diventa loro possibile, sulla base del metodo seguito, rinnovare radicalmente la gamma delle questioni anche di portata piú generale, su cui i pensatori si erano da sempre cimentati. Scoprono inoltre

di poter risolvere in certi casi in una maniera apparentemente definitiva quei problemi fondamentali, per i quali i vecchi metodi di pensiero si limitavano a indicazioni inconcludenti. È una scoperta che si è di continuo rinnovata con il passare degli anni e con il progredire della scienza. Se questo è vero, non deve stupire che lo scienziato sia rimasto fedele alle condizioni di partenza che hanno reso possibile la scienza. Anche se si occupa d'idee generali, resta essenzialmente uno specialista e continua a esprimersi come chi ritiene di poter ragionevolmente giungere a progressi decisivi soltanto attraverso la meticolosa accumulazione di milioni di fatti accertati con precisione (a cui la sintesi teorica segue).

La scienza che ha per oggetto il movimento dei corpi e le forze che ne sono causa o effetto, cioè la meccanica, viene creata da Galileo e da Newton. Nel corso del secolo XVII e all'inizio del XVIII essa si giova di due importantissime scoperte matematiche, quella della geometria analitica dovuta a Descartes, che permette di descrivere il movimento con un grafico e di ricondurne lo studio a quello delle curve, e soprattutto quella del calcolo differenziale e integrale, dovuta a Newton e a Leibniz, che permette di definire con precisione la nozione di velocità e d'introdurre quella di accelerazione. Si rende possibile per questa via una descrizione matematica della cinematica, scienza del movimento.

La dinamica, come scienza delle forze implicate nel moto, si edifica simultaneamente intorno al principio d'inerzia, dovuto a Galileo, e intorno alle leggi di Newton. Il principio d'inerzia afferma che in assenza di una qualsiasi forza un oggetto descrive una traiettoria rettilinea a velocità costante. Piú precisamente, esso afferma l'esistenza di sistemi di riferimento o «sistemi di assi» privilegiati, detti sistemi inerziali, in cui la situazione è proprio questa. Le leggi di Newton riguardano il caso piú generale in cui l'oggetto in movimento è sottoposto a forze. Esse affermano che in tal caso l'accelerazione del mobile è proporzionale alla risultante delle forze.

È risaputo che a Newton è inoltre dovuta la prima – senza dubbio, la piú geniale – applicazione delle leggi della dinamica. Partendo dalla sola ipotesi dell'esistenza tra due corpi celesti distanziati di forze attive a distanza (dette forze «gravitazionali»), dirette lungo la retta che congiunge i due corpi, proporzionali al prodotto delle loro masse e inversamente proporzionali al quadrato della distanza, Newton riuscí difatti a mostrare che ciascuno di questi corpi percorre necessariamente un'ellisse, che ha uno dei fuochi nel centro di gravità del sistema. Nel caso del Sole e di un suo qualsiasi pianeta, il rapporto delle masse dei due corpi fa sí che questo fuoco coincida all'incirca con il Sole. La teoria mostra in questo modo che i rapporti regolari scoperti da Keplero nelle reciproche relazioni tra le traiettorie dei pianeti e nel movimento di ciascuno di questi, non sono altro che conseguenze del sistema di attrazione universale appena descritto (tra l'altro, è possibile notare che questa attrazione spiega anche il peso e il moto della Luna).

È difficile sopravvalutare la portata delle ricerche di Copernico, Galileo e Newton. È grazie a loro che l'uomo scopre contemporaneamente di essere

solo un atomo sperduto nell'immensità dell'Universo, ma di essere anche capace di comprendere questo Universo. Val dunque la pena di analizzare brevemente i principî di un metodo cosí fecondo di risultati.

Lo si può fare nel modo migliore facendo riferimento alla dinamica delle idee di quell'epoca, che si organizza attorno a due poli opposti, da una parte l'atteggiamento empiristico, dall'altra quello razionalistico.

L'atteggiamento empiristico è relativamente nuovo. Come si è visto, data soltanto dal Rinascimento. È questo atteggiamento che spinge a cercare la verità con metodi basati essenzialmente sull'osservazione e sulla sperimentazione. Raccomandato con molta energia dai pensatori inglesi, è riassunto e filosoficamente fondato nella celebre formula «Non vi è nulla nell'intelletto che non sia passato prima per i sensi». La storia degli sviluppi ulteriori della scienza ha ampiamente mostrato la fecondità pratica di una formula tanto favorevole al rispetto dei fatti osservati. Da un punto di vista teorico va tuttavia notata una cosa che gli empiristi sovente dimenticano: anche se la formula citata è giusta, essa non prova che quanto si trova infine nello spirito sia conforme a una «realtà esterna», data per esistente indipendentemente dallo spirito umano.

L'atteggiamento razionalistico, da parte sua, fonda la propria ragion d'essere proprio su questo rilievo. Piú in generale, sulla legittimità, già notata in quanto precede, di un dubbio espresso con chiarezza da Descartes: posso ben osservare e misurare, ma se il mio fine ultimo è di conoscere la realtà con certezza, con questo metodo fallirò necessariamente, poiché le sole certezze che può eventualmente fornirmi riguardano esclusivamente l'apprendimento dei risultati di queste misure da parte della mia coscienza, e in nessun modo la Realtà (a meno che io non faccia l'ipotesi di una corrispondenza: ma è proprio la necessità di questa ipotesi indimostrabile che esclude ogni accesso alla certezza). È per questo che l'atteggiamento razionalistico raccomanda altre vie di accesso alla conoscenza dell'Essere: vie che fanno affidamento, secondo la tradizione platonica, sulla sola forza del pensiero.

Le scoperte di Galileo e di Newton, una volta che furono conosciute e confermate, contribuirono in larga misura a troncare la controversia, a scapito soprattutto dell'atteggiamento razionalistico. Descartes da parte sua fallí in pieno nel tentativo di edificare una fisica valida col suo metodo, basato su un razionalismo radicale. È, questa, una circostanza ricca d'insegnamenti: mostra infatti che il pensiero, anche intelligente, fuori dal suo ambito abituale che è di guida dell'azione si lascia fuorviare molto facilmente. Dimostra cosí che il pensiero può in particolare lasciarsi ingannare da ogni sorta d'ipotesi ingiustificate, che non si avvede nemmeno di porre. Se la storia della fisica portasse pur come solo contributo la scoperta di questa relativa debolezza del puro pensiero, e riuscisse davvero a farne comprendere gli autentici termini a quella categoria di persone che si professano pensatori, un risultato cosí importante giustificherebbe da solo che all'insegnamento della storia della fisica si dedichino grandissimi sforzi.

Ma, d'altra parte, sarebbe evidentemente illegittimo presentare i successi

di Galileo e di Newton come una lampante giustificazione, senza riserve, dell'atteggiamento radicalmente empiristico dei pensatori inglesi dell'epoca. Se la teoria di Newton trae tutto il suo valore dalla conferma che ne dà l'osservazione, non è meno vero che essa fa ricorso a una matematica sofisticata, la cui conoscenza da parte dell'uomo non è affatto evidente (ed è il meno che se ne possa dire) che provenga dai dati dei sensi.

La scoperta della gravitazione universale è in realtà un esempio eccellente del modo in cui la fisica è capace di combinare l'empirismo e il razionalismo invece di contrapporli. Come non ci sarebbero riusciti da soli i dati delle osservazioni, per quanto numerose, cosí nemmeno la matematica pura sarebbe stata in grado di fornire, essa sola, la conoscenza cercata. La loro unione è invece stata feconda perché, come si è visto, le osservazioni di Tycho poterono essere sintetizzate da Keplero in formule matematiche semplici, delle quali la geniale ipotesi di Newton rende conto in modo eccellente: cosa che quest'ultimo è riuscito a dimostrare, ma solo con l'impiego di tecniche matematiche elaborate. Ecco qui un compendio quasi perfetto di quel che è il metodo in fisica.

La conoscenza delle scoperte qui sopra ricordate – conoscenza che si diffuse nei secoli XVII, XVIII e all'inizio del XIX – rappresentò una fortissima motivazione alla ricerca scientifica, che si sviluppò in svariate direzioni. In particolare, l'ottica fece in questo periodo grandissimi progressi. Piú tardi apparve la chimica con la scoperta dei corpi puri: scoperta ben presto seguita da quella degli elementi chimici, che esistono in numero definito e ristretto e dànno origine a tutti i costituenti della materia combinandosi secondo proporzioni definite. Verso la fine del periodo apparve anche la scienza dell'elettricità, prima con l'elettrostatica – della quale Coulomb riuscí a dimostrare l'analogia con la gravitazione newtoniana – e poi con l'elettrodinamica, o scienza della corrente elettrica e delle sue principali proprietà (Volta).

Nello stesso tempo, la fecondità della teoria di Newton contribuí a spingere molti spiriti di quest'epoca – e non soltanto di quelli che si dedicavano alla scienza – a ritenere che la ragione, aiutata o meno dall'esperienza, può con relativa facilità aprirsi un accesso autentico alla totalità di ciò che *è*, alla sola condizione di liberarsi dei pregiudizi. È la filosofia del secolo dei lumi, che l'*Encyclopédie* di Diderot e d'Alembert illustra bene.

Da parte sua era stato però ben piú modesto Newton che soleva immaginarsi come un bambino sulla spiaggia intento a raccogliere per gioco i ciottoli colorati piú vivamente di altri, mentre l'immenso oceano del Reale gli si stendeva innanzi ancora vergine e inesplorato. Pur avendo introdotto lui stesso la nozione di forza agente a distanza, con il successo che si sa, vedeva ad esempio in tale nozione delle difficoltà, che sembrano averlo condotto, e giustamente, a ritenere che dovesse trattarsi solo di una tappa nell'approfondimento della conoscenza della realtà da parte dell'uomo.

Ma senza dubbio furono soprattutto alcuni filosofi dell'epoca che seppero nel miglior modo scoprire e mettere in luce aspetti troppo ingenui e schematici dell'ottimistica credenza in una facile conoscenza dell'essere attraverso i metodi dell'empirismo o della ragione. A tal riguardo le loro idee furono premonitrici per la fisica. Non è possibile, dunque, passarle qui sotto silenzio. In particolare, Hume nel secolo XVIII sottolineò correttamente sia il ruolo essenziale dell'induzione in qualsiasi conoscenza generale sia il suo carattere di postulato indimostrabile (dire che questo postulato è almeno verosimile è già compiere un'induzione, concernente l'idea di *probabile*).

Contemporaneamente, riemergeva la vecchia questione di Descartes: Berkeley, nella fattispecie, non aveva alcuna difficoltà a riscoprire il fatto che, se è vero che l'uomo non ha dell'essere una conoscenza diretta, se è vero d'altra parte che per acquisire una conoscenza deve osservare – e, se cerca la precisione, misurare –, allora i soli fatti che è in grado davvero di affermare sono le impressioni che ha quando compie queste operazioni. Pensare che gli altri uomini abbiano anch'essi delle impressioni e che queste, quando si tratti di esperienze banali, siano coerenti con le proprie, è già per il singolo individuo frutto di ipotesi; e lo è ancor piú il voler superare il «tutto accade come se» e pretendere dunque che la somma totale delle nostre osservazioni, pur essendo riassumibile in una teoria elegante come quella di Newton, descriva effettivamente qualcosa che esiste indipendentemente da noi e che esisterebbe anche in mancanza di qualsiasi osservatore. Berkeley cosí nega persino che il concetto stesso di una tale esistenza *abbia un senso*. Essere, afferma, è soltanto essere percepito.

È noto quel che venne dopo. In particolare, come Kant, colpito da tali argomenti e soprattutto dalla difficoltà sollevata da Hume circa l'induzione, fosse condotto a quella che egli stesso chiamò la sua «rivoluzione copernicana». Per darne un rapido schema, va ricordato che secondo Kant l'essere – o, come lui lo chiama, la cosa in sé – è rigorosamente inconoscibile. Come uno, nato con degli occhiali blu che non può togliersi, vede il mondo di un unico colore, pur essendo questo variamente colorato, cosí l'uomo, nato con certe strutture a priori che sono proprie della sensibilità umana, percepisce l'essere attraverso queste stesse strutture, o, per meglio dire, le impone lui stesso alla totalità della sua apprensione. È in questo modo che la nozione di causa, in particolare, viene considerata da Kant come riferentesi essenzialmente alla struttura dello spirito umano che la proietta nella propria visione delle cose: il che aiuta a capire il ricorso all'induzione. Ma, in modo piú generale, Kant pensa che lo stesso accada per nozioni fondamentali come quelle di spazio e tempo, che egli vede non come entità esistenti in sé (o proprietà di entità esistenti in sé), ma come condizioni a priori della sensibilità propria dell'uomo.

Se la cosa in sé è inconoscibile, è inutile allora parlarne. Ma in questo modo le parole 'la Realtà', 'il Reale', diventano parole disponibili per un uso diverso. Visto che l'esperienza umana, in particolare l'esperienza scientifica, presenta una certa regolarità (vi si possono scoprire delle leggi), è possibile ancora continuare a dire che essa descrive il Reale – la Realtà – a patto di cambiare il significato dei termini. E infatti i kantiani usano spesso la parola 'realtà'. Ma bisogna tener presente che nelle loro intenzioni questa parola

non si riferisce a una realtà indipendente dall'uomo, ma piuttosto a una realtà empirica, che riflette almeno in parte la struttura dello spirito umano.

2.3. Secolo XIX.

Con l'inizio del secolo XIX il numero sempre crescente delle conoscenze rende inevitabili le specializzazioni e, in particolare, scienziati e filosofi sviluppano ormai le proprie ricerche di solito separati gli uni dagli altri. I primi non si curano piú molto dei lavori dei secondi, i quali, il piú delle volte, rendono loro la pariglia. Sarebbe dunque inutile presentare qui un rendiconto, anche succinto, dello sviluppo della filosofia da quel momento ai giorni nostri. Tuttavia la separazione non fu del tutto rigorosa. Da una parte, infatti, si costituí l'epistemologia, o filosofia della scienza. Dall'altra, alcuni fisici alle prese con apparenti paradossi o comunque con difficoltà d'interpretazione dei risultati che sembravano loro particolarmente rilevanti pensarono di potersi utilmente riferire alle riflessioni e alle opere dei filosofi per arrivare a una soluzione. Fu questo il caso, nel secolo XX, di Heisenberg che, per interpretare la fisica quantistica, mutuò esplicitamente importanti elementi proprio dal pensiero di Kant, come si avrà modo di vedere nel § 4.

Le basi dell'attuale dominio tecnologico dell'uomo furono fondate nel corso della prima metà del secolo XIX con la comparsa della termodinamica (Carnot, piú tardi Joule e Clausius) e con la scoperta delle leggi dell'elettrodinamica (Biot e Savart, Laplace, Ampère). È importante notare che questi sviluppi poterono aver luogo senza che i loro autori conoscessero la natura «reale» dei fenomeni che studiavano: fu solo piú tardi, cioè con la scoperta delle leggi essenziali della termodinamica e dell'elettrodinamica, che l'agitazione molecolare fu riconosciuta come causa dei fenomeni calorici e che divenne noto il carattere granulare dell'elettricità (esistenza degli elettroni). Questo prova a sufficienza che per compiere dei progressi significativi – talvolta anche di grande portata – lo scienziato non ha necessariamente bisogno di conoscere «la natura ultima delle cose». Si tratta di un fatto estremamente importante, che da un certo punto di vista risulta positivo, poiché tale natura ultima – i filosofi da duemila anni ne fanno l'amara esperienza! – è ben difficile da conoscere. È dunque una fortunata combinazione che anche nei campi che non hanno nulla di filosofico un tal sapere non si riveli necessario per l'elaborazione di un sistema di conoscenze, che sia insieme strutturato e fecondo e che prenda l'aspetto di una teoria fondamentale che comporta sistemi di equazioni, formulazioni precise dei risultati dell'osservazione derivanti da condizioni iniziali date e cosí via. Da un altro punto di vista, il fatto empirico or ora constatato – e cioè che la conoscenza della natura ultima delle cose non è in realtà *necessaria* per l'elaborazione di una teoria scientifica – deve evidentemente rendere lo scienziato prudente quando abbandona esplicitamente il campo dei fenomeni e pretende di descrivere *ciò che è*.

Quest'ultima considerazione va raffrontata col fatto, ben noto ai fisici teorici, che molto spesso fenomeni assai differenti tra di loro – persino appartenenti a discipline diverse – obbediscono alle stesse equazioni. Tali similarità tra differenti discipline, dovute alle corrispondenze della loro struttura matematica, possono in alcuni casi risultare meramente accidentali. In altri casi possono invece rivelare autentiche identità, nascoste sotto un'apparenza di diversità. Di fatto una delle caratteristiche essenziali dello sviluppo della fisica nel corso del secolo XIX fu proprio la sua fecondità nel campo dell'unificazione – attraverso la matematica – di discipline le cui reciproche relazioni erano fino a quel momento rimaste nascoste. È cosí che con l'avvento della meccanica statistica (Boltzmann, Gibbs) l'intera termodinamica poté essere collegata alla meccanica; è cosí che, per essere piú precisi, la termodinamica fu allora costituita come sintesi originale, ricca di problemi filosofici, della meccanica e del calcolo delle probabilità, calcolo che avevano contribuito a creare matematici come Pascal, Eulero e i Bernoulli. È ancora in questo modo che le discipline prima distinte dell'elettrodinamica e del magnetismo si fusero in una sola – l'elettromagnetismo – grazie alle scoperte di Faraday. E soprattutto è ancora cosí che in una successiva fusione, ancora piú importante della precedente, l'elettromagnetismo giunse a inglobare l'ottica.

Esposto in modo molto schematico, il nucleo centrale dell'elettromagnetismo può essere descritto come fondato sulla scoperta del fatto che le cariche elettriche che la materia può portare (cariche quantizzate, positive o negative) sono sottoposte a due tipi di forze, che vengono rappresentate con dei campi, cioè schematicamente con il valore della forza – elettrica ad esempio – esercitata su una unità di carica in ciascun punto dello spazio. Nel caso statico (caso in cui nulla si muove) le forze in questione sono beninteso esse stesse create da cariche situate in altri luoghi, che hanno nei loro confronti il ruolo di «sorgenti», di modo che le forze potrebbero essere considerate come proprietà a distanza di queste cariche. Tuttavia sotto l'impulso di Faraday i fisici furono portati progressivamente a considerare la possibilità di conferire mentalmente ai campi corrispondenti una sorta di esistenza locale, in qualche modo indipendente dalle loro «sorgenti». Si scoprí allora che il campo elettrico e quello magnetico sono legati l'uno all'altro da leggi semplici, in particolare quando variano nel tempo. Verso il 1860 le leggi furono sintetizzate da Maxwell in un insieme di equazioni che porta il suo nome. Un'importante conseguenza matematica di queste equazioni è la predizione dell'esistenza di oscillazioni che si propagano sotto forma di onde e che non sono nient'altro che le onde elettromagnetiche (Hertz, Marconi) utilizzate in seguito dalle tecniche della radiofonia e della televisione. La loro velocità è quella della luce e infatti, dopo la loro scoperta, si poté stabilire con certezza che la luce non è altro che un'onda elettromagnetica: scaturiva di qui la sintesi dell'ottica e dell'elettromagnetismo.

La seconda metà del secolo XIX è un'epoca di progressi notevolissimi nella conoscenza della struttura della materia. L'esistenza di atomi e molecole, già resa verosimile dagli sviluppi della chimica, è confermata da Jean Perrin, che dimostra come i loro movimenti casuali dovuti all'agitazione molecolare sono all'origine del moto browniano, cioè dell'agitazione che si può osservare in

piccolissimi oggetti posti in un liquido (anche se molto grossi in rapporto agli atomi o alle molecole, questi oggetti risultano abbastanza leggeri per essere messi in movimento quando un atomo o una molecola entra in collisione con loro). D'altra parte l'invenzione di apparecchi elettrici ad alta tensione rivela l'esistenza delle prime particelle elementari, il protone e l'elettrone. Quest'ultima particella viene riconosciuta all'origine dell'esistenza della corrente elettrica, che risulta costituita dallo spostamento di un gran numero di tali elettroni nei fili conduttori.

In questa stessa epoca vengono compiuti grandi progressi anche per quel che riguarda la nozione di energia. Ai concetti di energia meccanica, cinetica e potenziale erano già venuti ad aggiungersi quelli di energia termica, elettrica e chimica. Nello stesso tempo la legge di conservazione dell'energia si era rivelata universale. Ora viene a precisarsi la profonda unità di queste diverse forme di energia. In particolare, l'energia termica è ora concepita come essenzialmente di natura meccanica e piú precisamente cinetica, in quanto dovuta all'agitazione degli atomi e delle molecole.

La relazione della fisica con la situazione generale delle idee nella seconda metà del secolo XIX può essere schematizzata nel modo seguente. Una gran parte del mondo dei filosofi, per influenza delle idee di Kant, o ancor piú di quelle di Hegel e dei suoi successori, si tiene con molta prudenza lontano da sviluppi che sembrano insieme troppo astratti, troppo tecnici e – paradossalmente – forse troppo ingenui. Troppo ingenui, almeno nel senso che la maggioranza dei fisici dell'epoca sembra far affidamento sul concetto di realtà in modo assai poco critico. Dall'altra parte, in larga maggioranza, i fisici, incoraggiati dai propri successi, non vedono perché dovrebbero porsi a proposito di un concetto cosí generale dei problemi filosofici che sembrano loro falsi problemi.

Si sviluppano invece discussioni sui metodi: dando il cambio agli enciclopedisti, Auguste Comte continua la loro battaglia contro i «pregiudizi», teisti o metafisici, e insiste, per antitesi, sull'imperativo che prescrive di accordare senso solo ai fatti e alle teorie che collegano i fatti gli uni con gli altri. Getta cosí le basi del metodo positivistico. Peirce e James, i fondatori del pragmatismo, in modo molto simile sottolineano l'importanza del criterio di utilità (che può d'altronde non avere di mira che la dimensione teorica) per giudicare del valore di una teoria o concetto. La critica mossa da Mach alle idee newtoniane di forza e di massa va in certa misura nella stessa direzione.

Comte sembra da parte sua accettare in modo implicito, e quindi non critico, l'ipotesi che un fatto – cioè secondo il suo metodo, un insieme di dati osservati – è necessariamente un elemento di una realtà totalmente indipendente dall'uomo. Egli non sembra in effetti essersi posto la questione centrale del problema della conoscenza, che era stata in precedenza sollevata da Descartes o da Berkeley. In una certa misura, si potrebbe riscontrare un'analoga lacuna nei tre pensatori citati e in parecchi dei loro emuli, che non rilevano, come si dovrebbe, l'ambiguità della parola 'fatto' nel contesto delle loro dottrine. Resta comunque vero che l'elogio dei metodi scientifici fatto dai

positivisti, congiuntamente coi primi risultati tecnologici ottenuti con tali metodi, ebbe per conseguenza verso la fine del secolo XIX un grande entusiasmo popolare per le scienze in generale e, in particolare, per la piú perfezionata di esse, la fisica. Anche eminenti letterati contribuirono in questo periodo a volgarizzare l'idea di una scienza che sarebbe da una parte in grado di portare la felicità agli uomini grazie ai progressi tecnici che può procurare, e dall'altra di fornire la conoscenza ultima dell'essere, riducendolo agli atomi e ai campi, visti come costituenti di una realtà assoluta, che è completamente indipendente dall'uomo, anzi lo ingloba totalmente.

Va da sé che visioni tanto audaci, e di carattere cosí nettamente ontologico, non furono opera dei fisici e nemmeno degli epistemologi dell'epoca. Ma, d'altro lato, il positivismo e le sue varianti ebbero a partire da quest'epoca un'influenza diretta e, essa almeno, del tutto autentica sulla ricerca scientifica; e questo per ragioni di carattere metodologico. D'altronde tale influenza non ebbe sempre conseguenze felici. In Francia, ad esempio, il chimico Berthelot si oppose a lungo e con successo all'introduzione del concetto di atomo e al suo uso in chimica, proprio perché allora l'esistenza «reale» degli atomi non era ancora stata verificata e sembrava inverificabile. Il criterio positivista secondo cui una nozione non ha senso, a meno che non la si possa definire in relazione a un'operazione e a una misura, non risultava soddisfatto dalla nozione di atomo: di qui il suo rifiuto.

2.4. Secolo XX.

Tuttavia, in altre circostanze, lo stesso criterio positivistico fu per la fisica di notevole fecondità. L'evento rappresentato dalla critica da parte di Einstein della nozione di tempo universale ne fornisce un chiaro esempio.

All'origine di tale critica sta la constatazione (tipica a quell'epoca di ogni fisico) del fatto che l'esistenza delle onde elettromagnetiche manifestamente richiede quella di un supporto, cioè di «qualcosa» che oscilli, e che non potrebbe essere alcuno dei mezzi già conosciuti, poiché la propagazione della luce negli spazi interstellari mostra immediatamente che tale mezzo dovrebbe essere universale. Rispetto a questo mezzo universale le onde elettromagnetiche hanno per le equazioni di Maxwell una velocità ben determinata, mentre ci si potrebbe «ingenuamente» aspettare che, in rapporto a un oggetto qualsiasi a sua volta in movimento nel mezzo in questione, abbiano una velocità (detta «relativa») diversa, esattamente come un'onda ha una velocità diversa rispetto all'oceano o alla riva, e rispetto a chi la segue facendo del surf. Ora, la struttura matematica delle equazioni di Maxwell lasciava sospettare che le cose non andassero proprio in questo modo: la fisica sperimentale riuscí a fornire in proposito un contributo decisivo. Il celebre esperimento di Michelson e di Morley permette di stabilire il fatto, in apparenza paradossale, che la luce viaggia sempre alla medesima velocità (300 000 Km/s) in rapporto a un oggetto qualsiasi, qualunque sia il modo in cui quest'ultimo si muove rispetto a qualsiasi altro oggetto.

I matematici dell'epoca che, come Poincaré, s'interessavano di fisica e pure alcuni fisici orientati, come Lorentz, verso la teoria, portarono a termine con successo uno studio matematico estremamente completo delle proprietà delle equazioni di Maxwell, esaminate dal punto di vista degli effetti osservabili, che queste equazioni potevano avere per degli osservatori in movimento relativo gli uni rispetto agli altri. Talvolta si è detto addirittura che sono stati questi due autori i veri creatori, dal punto di vista matematico, della teoria che doveva poi essere la relatività. Ma la loro costruzione teorica mancava di un elemento essenziale. Questa lacuna aveva come conseguenza il fatto che la spiegazione del «paradosso» descritto poco sopra comportava la necessità di postulare taluni fenomeni di contrazione delle lunghezze o di ritardo degli orologi e che, nella concezione tradizionale del tempo e dello spazio propria di questi autori, fosse insieme *indispensabile* e *impossibile* giustificare fisicamente questi fenomeni.

È a questo punto che intervenne la critica einsteiniana delle nozioni di spazio assoluto, di tempo universale e di simultaneità in sé. Con un'analisi serrata delle possibilità di misura, anche teoriche, di cui l'uomo dispone, Einstein poté stabilire che queste diverse nozioni non sono – come si dice oggi – affatto suscettibili di definizioni *operative*. Di conseguenza nulla ci obbliga a pensare che esse necessariamente corrispondano a proprietà della realtà indipendente. A fortiori, nulla nemmeno ci costringe a considerarle, come voleva Kant, come condizioni a priori intangibili e supreme, di ogni conoscenza umana. Perciò i problemi legati alla riflessione sullo spazio e sul tempo apparvero ad Einstein come problemi fisici paragonabili agli altri, per la cui soluzione era cioè opportuno *costruire* le nozioni piú idonee alla descrizione dei fatti. Come dimostra lo studio delle onde elettromagnetiche, queste nozioni non sono né quella di spazio assoluto, né quella di tempo universale; occorrono altre nozioni, tra cui quella di spazio-tempo è la piú importante. La cosiddetta teoria della relatività ristretta, costruita in tal modo da Einstein, non sopprime i fenomeni di contrazione delle lunghezze e di ritardo degli orologi, ma fa di meglio: incorpora intimamente nella teoria questi fenomeni, indispensabili per la spiegazione dei fatti osservati, in modo che lungi dall'apparirvi come inspiegabili, essi vi assumono l'aspetto di conseguenze naturali. Inoltre questa teoria conduce a modificazioni della meccanica apprezzabili alle grandi velocità e i cui effetti sono osservabili e osservati.

Dal punto di vista matematico il formalismo della relatività ristretta rivela che, non diversamente dalle leggi della meccanica, le equazioni di Maxwell restano immutate in tutta una classe di sistemi di riferimento, detti galileiani, dove la velocità della luce resta peraltro la stessa. Ma perché questa limitazione a particolari sistemi di riferimento? Questo grave problema unito ad alcuni altri (come ad esempio quello dell'incompatibilità tra la relatività ristretta e l'azione istantanea a distanza di Newton) portò Einstein a riprendere la sua teoria e a costituire la relatività generale. La caratteristica essenziale di questa teoria consiste nell'introdurre l'idea secondo cui lo spazio-tempo possiede una curvatura e nel rendere conto degli effetti di gravitazione

per mezzo di una tale curvatura. Se due oggetti pesanti tendono ad avvicinarsi l'uno all'altro, questo non accade in virtú di una forza di gravitazione attrattiva – come pensava Newton – ma, dice Einstein, semplicemente per una curvatura dello spazio, essa stessa generata dalle masse di questi oggetti. Sotto il profilo quantitativo si ritrovano gli effetti predetti dalla teoria newtoniana, ma il punto di vista è come si vede molto diverso. Inoltre la relatività generale predice anche alcuni effetti che le sono specifici e la cui esistenza è stata empiricamente verificata.

L'uso da parte di Einstein di argomenti operazionistici (come si direbbe oggi) nella sua critica delle concezioni classiche dello spazio e del tempo, fece pensare ai fisici positivisti suoi contemporanei che egli fosse in realtà uno di loro. Niente di piú falso, almeno se per «positivismo» s'intende la concezione radicale che fu propria del «circolo di Vienna» e che condusse i suoi sostenitori a respingere, come priva di senso, la nozione stessa di una realtà indipendente dallo spirito umano. Einstein, in effetti, era un realista della schiatta di Spinoza e di Leibniz. Ciò significa da una parte che il modo meno inverosimile di spiegarsi la regolarità della nostra esperienza, come lo stesso accordo intersoggettivo circa i fatti, gli sembrava quello di attribuirle, come «tutti» fanno, all'esistenza di una qualche realtà indipendente e conoscibile; e dall'altra che egli però ammetteva pienamente che tale realtà potesse avere una struttura molto diversa da quella attribuitale dalle forme di pensiero intuitive o classiche, essendo scontato che l'uomo aveva modellato queste ultime sulla capacità di azione che aveva un tempo, che concerneva solo i corpi a sua misura ed era pertanto assai limitata.

Agli inizi del secolo XX era legittimo pensare che i progressi della fisica sarebbero continuati ad andare nel senso di una conferma delle prospettive generali di Einstein, quasi senza che ci fosse il rischio d'imbattersi in ostacoli capaci di far dubitare del giusto fondamento di queste ultime. Piú in dettaglio: perfino i tentativi di una radicale geometrizzazione di tutta la fisica sembrarono a un certo punto destinati al successo. Ma poi ci si dovette arrendere all'evidenza. Se effettivamente la forza di gravitazione è geometrizzabile, lo stesso non vale per altre forze. La massima «Tutto è geometria» non può piú essere incisa sui frontoni degli istituti di fisica matematica.

Ma c'è qualcosa di ben piú grave. A partire dagli anni '20 non viene solo rimesso in causa l'ideale «geometrico» di Einstein, ma anche la sua piú generale concezione «realista». Questa evoluzione è dovuta alla comparsa della meccanica quantistica, ai suoi successi sbalorditivi e alle difficoltà d'interpretarla altrimenti che in una prospettiva fortemente tinta di positivismo radicale.

Ovviamente non si riporta nei dettagli lo sviluppo di questa teoria: si ricordi semplicemente che le onde elettromagnetiche si rivelarono (Planck, Einstein) dotate sotto certi aspetti di proprietà corpuscolari e che, simmetricamente (Louis de Broglie), le particelle (elettrone, protone, ecc.) si rivelarono dotate di talune proprietà ondulatorie. Ma come si poteva concepire una sintesi del genere tra proprietà apparentemente contraddittorie? Come si può facilmente indovinare, è possibile ottenere un tale risultato solo rinunziando

a un gran numero di intuizioni. Un costo cosí alto viene però controbilanciato, per fortuna da una gran messe di risultati, sia teorici sia applicati, relativi alla composizione e alle proprietà della materia.

La meccanica quantistica, come si presenta oggi, è opera soprattutto della scuola di Copenhagen (Bohr, Heisenberg, Jordan), di Dirac, Born, Schrödinger e Louis de Broglie. Essa si fonda sulle seguenti idee.

Nella misura del possibile conviene evitare di parlare di proprietà possedute da un sistema fisico, in altri termini di ricorrere a proprietà che un sistema fisico avrebbe «in sé». Di fatto, prima di averle misurate, noi ignoriamo queste proprietà e ne ignoriamo perfino l'esistenza (proprio come ignoro se la persona che risponde «sí» a una domanda che pongo abbia una precisa opinione sull'argomento prima che la domanda venga posta, o se risponda a caso). Dopo la loro misurazione quel che si conosce con certezza è soltanto il risultato della misura in questione. È quindi solo sui risultati delle diverse misure possibili che la meccanica quantistica formula delle leggi.

Queste leggi sono in generale di natura probabilistica. Ciò contrasta con il determinismo rigoroso postulato dalla fisica classica.

La probabilità del risultato di una misura di una certa grandezza fisica su un certo sistema in un certo istante è ottenuta componendo in modo adeguato un'entità matematica che descrive la grandezza e un'altra entità, detta «funzione d'onda» o «vettore di stato», che descrive il sistema fisico nell'istante dato. La descrizione di un sistema fisico in termini di funzione d'onda sostituisce la descrizione usuale in termini di posizioni e velocità delle particelle del sistema. Ma tale descrizione non è sempre possibile. Ci sono infatti casi in cui a un sistema dato non si può attribuire alcuna funzione d'onda. Questo succede, a rigore, con i sistemi che hanno interagito con altri sistemi i quali non possono essi stessi essere descritti nella maniera classica (cioè con la fisica precedente), poiché in questo caso risulta definibile soltanto la funzione d'onda dell'insieme. Non si potrà dunque fare a meno d'introdurre nei calcoli la funzione d'onda – inutilizzabile perché troppo complessa! – dell'intero universo dei sistemi fisici che hanno interagito con quello che ci si propone di studiare? Per un realista questa è la fonte di una difficoltà concettuale che si rivela autentica a uno studio piú approfondito. Fortunatamente tale difficoltà resta puramente di principio e di conseguenza ogni spirito «positivo» o «pragmatico» può non tenerne conto. In pratica i dispositivi di preparazione dei sistemi fisici che la meccanica quantistica si propone di studiare sono dispositivi materiali che costituiscono essi stessi dei sistemi fisici descrivibili in maniera «classica», cioè, ancora una volta, per mezzo dei concetti della vecchia fisica. I dispositivi di cui si tratta costituiscono quindi un dato di riferimento adeguato per la definizione non ambigua della funzione d'onda – o delle funzioni d'onda possibili – del tipo di sistema del quale ci si occupa.

L'evoluzione nel tempo della funzione d'onda ubbidisce a un'equazione alle derivate parziali detta «equazione di Schrödinger», che sostituisce le usuali equazioni della meccanica. Associata ad opportune condizioni al contorno questa stessa equazione permette di determinare quali sono le possibili funzioni d'onda di un sistema fisico, le cui proprietà osservabili non variano nel corso del tempo (come nel caso di una molecola stabile o di un atomo non eccitato).

Una conseguenza della teoria è dunque che i risultati delle misure di certe grandezze fisiche non possono prendere che certi precisi valori ben distinti gli uni dagli altri (valori quantizzati). Questo è il caso ad esempio dell'energia intrinseca che può avere un atomo o una molecola. A ciascun valore dell'energia possono poi corrispondere sia una funzione d'onda sia piú funzioni d'onda possibili.

La funzione d'onda di una particella specifica in generale solo la probabilità di trovare la particella nel tale o nel tal altro punto, nel caso di una eventuale misura della sua posizione: spesso essa permette dunque di dire con una quasi-certezza solo che la particella si trova in una *zona* ben determinata. Sia Δx lo spessore di questa zona. Analogamente la funzione d'onda in questione specifica in generale solo una «zona» simile di valori, di spessore, ad esempio, Δp_x, relativa a un'eventuale misura dell'impulso. È una conseguenza della teoria che Δx e Δp_x siano legate dalla relazione $\Delta x \cdot \Delta p_x > \hbar/2$, detta relazione di Heisenberg (o «relazione d'indeterminazione» anche se la parola risulta equivoca), e nella quale $\hbar = h/2\pi$ è legato a una costante universale, h, introdotta da Planck.

La meccanica quantistica è stata lo strumento teorico di base per i progressi della fisica negli ultimi cinquant'anni. I risultati piú notevoli sono i seguenti.

Onde di materia. Prima sintesi concettuale materia-luce. Invenzione della microscopia elettronica.

Fisica atomica e molecolare. Calcolo dei livelli di energia degli atomi e delle molecole. Di qui, per semplice differenza, determinazione delle frequenze della radiazione elettromagnetica che gli atomi e le molecole possono emettere e che possono assorbire. In particolare, la matassa (di incredibile complessità) degli spettri di radiazione degli elementi chimici ha potuto in tal modo essere interamente dipanata dalla teoria, a partire dai dati di base piú semplici, come massa e carica dell'elettrone, costante di Planck e velocità della luce.

Calcolo delle proprietà chimiche dei corpi, cioè spiegazione, almeno teorica, dei fatti che costituiscono la chimica. Qui la limitazione è costituita solo dalla complessità dei calcoli, che cresce molto rapidamente quando ci si allontana dai corpi di piú semplice composizione chimica. Il principio di una tale chimica teorica risiede semplicemente nella soluzione dell'equazione di Schrödinger per un sistema costituito da nuclei atomici dati e da un numero dato di elettroni. Le possibili energie cosí ottenute dànno delle informazioni sulla stabilità del complesso in tal modo definito e le funzioni d'onda corrispondenti informano sulle strutture possibili di quest'ultimo (forma delle molecole nello spazio, ecc.).

Fisica dello stato solido. Spiegazione delle proprietà elettriche degli isolanti, dei conduttori, dei semiconduttori. Spiegazione della superconduttività. Calcolo delle principali proprietà dei solidi. Applicazioni ai semiconduttori, invenzione dei transistor e dei laser.

Fisica nucleare. Analisi dettagliata della struttura dei nuclei atomici. Applicazione alla fissione e alla fusione, energia nucleare.

Elettrodinamica quantistica e teoria quantistica dei campi. Accertamento delle proprietà del fotone o particella della radiazione elettromagnetica. Predizione dell'esistenza dell'antimateria. Sintesi concettuale materia-radiazione. Studio delle particelle recentemente scoperte e classificazione delle loro interazioni. Analisi della struttura dei protoni e dei neutroni. Formulazione dell'ipotesi dei quark, particelle di carica frazionaria rispetto a quelle dell'elettrone e del protone, i quali – in gruppi legati da campi detti «colla» – costituirebbero i protoni e i neutroni a loro volta costituenti i nuclei. Formulazione (in corso) di una teoria coerente di queste strutture.

La meccanica quantistica e i suoi prolungamenti hanno ampliato immensamente l'orizzonte dell'insieme delle scienze esatte. Ma, ancora una volta, questo è accaduto a prezzo di una revisione talvolta lacerante di idee intuitive, ma fondamentali. Qualche volta si deplora la necessità di una tale revisione, ma invano: è meglio scorgervi la fonte possibile di un progresso in profondità nella comprensione delle relazioni che possono esistere tra lo spirito umano e la natura. L'ultimo paragrafo di questo articolo (§ 4) è dedicato proprio alla descrizione qualitativa dei principali problemi che attualmente si pongono in questo campo.

3. *Aspetti caratteristici della situazione della fisica nella scienza.*

3.1. Unità della scienza.

La parola 'scienza' ha oggi un'accezione molto ampia, poiché molte discipline che un tempo rientravano in quelle umanistiche sono giunte a rivendicarla. Si userà qui il termine nell'accezione «insieme di scienze empiriche dette esatte», con l'esclusione dunque delle cosiddette scienze umane come, ad esempio, la sociologia.

La scienza cosí intesa non è nient'altro che una giustapposizione di discipline a compartimenti stagni, che non hanno influenza le une sulle altre, o possiede invece un'unità? E, nella seconda ipotesi, qual è la natura di questa unità? Questo è il problema che si vuole affrontare qui alla luce delle conoscenze fornite dalla fisica contemporanea.

Per due almeno dei suoi caratteri la scienza presenta un'innegabile varietà. Uno di questi è esterno e dunque contingente. Riguarda il fatto che nella scienza e nella tecnologia la «visione d'insieme», la comprensione delle grandi idee fondamentali non basta. I dettagli sono fissi, esatti e devono essere quindi conosciuti. La loro moltitudine dà dunque luogo a un rigoglio di specialità di ogni sorta, ciascuna delle quali ha il proprio oggetto e i propri metodi particolari. L'insegnamento superiore stesso, che forma i futuri specialisti, incalzato da considerazioni utilitaristiche e dalla necessità di non allungare indebitamente la durata degli studi, si trova costretto a sacrificare una buona parte della cultura scientifica di base agli imperativi della specializzazione. Il risultato è una polverizzazione sempre maggiore delle conoscenze.

Sul piano dell'azione immediata questa polverizzazione quasi non presenta inconvenienti. Un informatico può far uso sempre piú di «scatole nere» senza preoccuparsi del loro contenuto. Può anche ignorare i termini delle recenti scoperte nella fisica di base che hanno reso possibile l'invenzione dei dispositivi che quelle scatole contengono. Analogamente il chimico che studia molecole può misurare tempi di reazione o energie di legame, invece di domandare al fisico di calcolarli esplicitamente con i metodi teorici della meccanica quantistica. È pur vero che il piú delle volte questo è dal suo punto di vista senz'altro preferibile, tenuto conto che, anche se fossero effettuati con calcolatori, tali calcoli risulterebbero in generale estremamente lunghi e, per tale ragione, praticamente irrealizzabili. Le relazioni tra la microbiologia e la chimica sono in parte dello stesso tipo. Esempi del genere potrebbero essere moltiplicati a volontà; la loro abbondanza mostra come difficoltà *pratiche* allo stato attuale ancora insuperabili finiscano per creare tra le discipline differenze qualitative e quasi delle barriere.

L'altra caratteristica della scienza che la rende cosí diversificata risiede in una certa carenza di interconnessioni dei modelli tra di loro, carenza che va riconosciuta, perché esiste *nei fatti* e non soltanto nelle lacune poco sopra ricordate della cultura scientifica generale degli specialisti. A questo riguardo la scienza ricorda un po' certi edifici che visti da lontano sono pieni di maestà e di armonia, ma che rivelano, a chi si avvicina, squilibri, parti in ricostruzione, rischiose impalcature. Anche all'interno di una data disciplina accade spesso, se si guarda da vicino, di vedere in contrasto concezioni e modelli stranamente differenti. Queste macchie che compaiono sul frontone compatto della scienza sono dovute alla sua incompiutezza come al fatto che è una tipica opera artigianale, costruita giorno per giorno da spiriti curiosi e dunque originali. Esse non devono preoccupare.

Infine, barriere tra discipline, difetti di interconnessione nei concetti e nei modelli, tutte queste fonti di diversità di punti di vista dipendono in parte dalla seguente circostanza. Se due scienze riguardano scale di dimensioni molto diverse, quella che ha per oggetto le strutture piú estese non può in generale far uso delle informazioni fornite dall'altra, se non in modo globale. Cosí, per limitarsi ad un esempio, lo specialista dei cristalli, o quello delle grandi molecole, pur avendo a che fare con degli aggregati di microggetti, deve ben guardarsi dall'usare tali e quali le equazioni di base che regolano la fisica dei microggetti. Sotto questo aspetto il numero delle variabili in gioco non per-

mette alcuna speranza. Lo specialista è dunque costretto a giocare d'astuzia: ricorre dunque a quei veri e propri relè che sono le costanti fisiche macroscopiche. I numeri che gli servono come dati di partenza sono in realtà calcolabili, o di fatto o almeno di diritto, mediante le equazioni della meccanica quantistica. Ma, piú semplicemente, possono essere misurati. Essi rappresentano i soli legami che uniscono questa meccanica alla disciplina macroscopica considerata, che per il resto si fonda su una concettualizzazione radicalmente diversa (quella della fisica classica).

Pur significative, queste sfumature non intaccano la fondamentale unità della scienza. Questa, incontestabile, è assicurata dalla fisica e chiunque si curi del significato della fisica deve prenderne coscienza. L'universo è immenso, ma ovunque ubbidisce alle stesse leggi fondamentali. Ecco il primo fatto incontrovertibile che bisogna sempre aver davanti agli occhi. Una galassia come la nostra non è che una molecola dell'universo: se ne sono contate altre, piú di cento milioni. Ma essa, da sola, contiene già alcune centinaia di miliardi di Soli. Noi vediamo corpi celesti la cui distanza da noi è dieci milioni di miliardi di volte superiore a quella che ci separa dalla Luna. E tuttavia siamo pressoché sicuri che in tutti questi immensi spazi le leggi che regolano i fenomeni sono le stesse a cui noi siamo soggetti e di cui le nostre teorie rendono conto.

Di fronte a questa sproporzione tra la debolezza dei nostri corpi e la potenza del nostro spirito già Pascal restava meravigliato. E infatti, che quel fuscello pensante che è l'uomo possa capire l'universo resta un enigma. È in parte per risolverlo che alcune concezioni epistemologiche, di cui si è vista la genesi, sostengono in via di principio che nei fatti e nelle leggi scientifiche l'uomo viene a conoscere soprattutto se stesso. Esse si fondano su argomenti del tutto solidi e richiederanno adeguata attenzione. Qui ci si deve soltanto premunire contro l'interpretazione superficiale di queste dottrine che sfocerebbe in una sorta di liquidazione della fisica, che verrebbe considerata di colpo una fantasia soggettiva o una collezione di ricette per la pratica. A questo proposito la cosa migliore che si possa fare è quella di tener presente la vasta accumulazione di osservazioni che dimostrano come fenomeni compresi ma complessi, conosciuti nella nostra regione del mondo, si producano anche altrove.

È quel che succede ad esempio con gli spettri luminosi degli atomi e delle molecole. È noto che ogni gas, portato ad incandescenza, emette luce. Analizzata con un prisma questa luce si divide in una moltitudine di *righe spettrali*. Le posizioni di queste ultime caratterizzano la composizione del gas. Infatti un gas formato da un corpo chimico elementare produce uno *spettro di righe* nel quale le distanze tra righe vicine sono perfettamente determinate e immutabili. Per di piú, questi valori numerici che ciascun elemento fornisce in numero grandissimo non sono soltanto dei fatti forniti dall'esperienza. Con la meccanica quantistica si dispone, come si è detto, di una teoria che permette di calcolarli, in modo esatto, a partire da nozioni del tutto generali. Ora, se si analizza in tal modo non piú la luce emessa da una qualunque fonte terrestre, ma la luce proveniente da una stella, o da una lontana galassia, si trovano proprio queste stesse righe, tra cui le distanze di cui si è detto sopra sono esattamente rispettate (a meno di un solo fattore correttivo: ma anche questo è ben capito e risulta utilissimo: misura la velocità detta «radiale» di questi oggetti lontani rispetto alla Terra). In questo modo è possibile avere la certezza che si trovano nei corpi celesti solo gli elementi disponibili sulla Terra e si riesce a calcolarne le proporzioni nell'universo.

Data la complessità di ciascuno spettro, è evidentemente impensabile, alla luce di queste osservazioni, una negazione pura e semplice dell'esistenza di leggi universali. Quando piú avanti si contesterà il carattere di verità assolute delle proposizioni interpretative della fisica, non si dovrà mai dimenticare che il loro carattere universale resta però accertato. E per bilanciare le tendenze allo scetticismo bisognerà tenere ben presente la conoscenza dell'estensione del mondo che solo la fisica fornisce. È chiaro che se, ad esempio, lo studio delle difficoltà dei fondamenti della fisica giungesse a farci assumere un atteggiamento puramente negativo nei suoi confronti, questo non potrebbe verificarsi che per mancanza di attenzione ai dati elementari, ma innegabili, or ora ricordati.

Ma anche noi dobbiamo ancora fare un passo avanti. Non è solo necessario inchinarsi davanti all'esistenza di leggi universali, ma bisogna anche ammirarne la coerenza. Anche questo è elementare, ma non deve essere dimenticato. L'ottica, ad esempio, è retta da proprie leggi, come quelle di Huygens o di Fresnel. L'elettromagnetismo è regolato da altre leggi, come quelle di Biot-Savart, di Ampère o di Faraday. E ovviamente le une non contraddicono le altre, perché si riferiscono a un altro oggetto. Ma c'è molto di piú. Si è visto che, a uno stadio ulteriore dello sviluppo delle discipline in questione, è stato possibile trovare delle leggi ancora piú generali (in particolare le equazioni di Maxwell) dalle quali derivano necessariamente *e* le leggi dell'ottica *e* quelle del magnetismo, come pure dell'elettricità. È in questo senso, molto forte, che bisogna qui intendere la parola 'coerenza'. Ed è in questo senso che oggi si può dire, almeno in prima approssimazione, che in realtà tutte le scienze della natura sono, almeno di diritto, coerenti le une con le altre. La comparsa della termodinamica statistica (con sfumature che saranno analizzate piú avanti) ha reso la termodinamica e la meccanica newtoniana coerenti tra di loro. In seguito è soprattutto la meccanica quantistica a rendere la chimica coerente con la totalità della fisica. Infine la teoria dei campi ha unificato, almeno formalmente, la fisica delle particelle e quella della radiazione.

Cosí al momento attuale l'aspetto della scienza che piú colpisce a prima vista è la sua notevole unità. Non si esagera molto dicendo che le equazioni di Maxwell, gli assiomi di base della meccanica quantistica e il principio di Pauli forniscono da soli la chiave dell'infinita varietà dei fenomeni, prodotti naturalmente, osservati sulla Terra. L'espressione «prodotti naturalmente» serve qui ad escludere certi fenomeni notevoli e ancora mal capiti, che si possono ad esempio osservare per mezzo dei grandi acceleratori. Questi fenomeni mettono in gioco forze e leggi diverse da quelle elencate sopra. Il fatto stesso che

per disporre di siffatti fenomeni sia necessario crearli artificialmente per mezzo di macchine onerose e complicate, ancor piú rinforza la nostra impressione di una straordinaria unità della scienza contemporanea. Si può descrivere questa impressione ricorrendo a un'immagine. Mentre una volta le varie scienze della natura erano come isole nel grande oceano dell'ignoto, sembra oggi che il paragone vada rovesciato. Dall'atomo alla stella, il campo delle nostre conoscenze forma un unico continente. Naturalmente la conoscenza dei fenomeni non può essere considerata completa. Restano dei laghi, per portare avanti il paragone, perfino dei mari interni, costituiti da fenomeni mal conosciuti o male spiegati. È cosí che ad esempio il magnetismo terrestre aspetta ancora una sua teoria definitiva. Ma, si sa, ciascuno di questi campi è limitato.

Quanto all'idea troppo diffusa secondo la quale i grandi progressi scientifici abbatterebbero questi o quegli aspetti della scienza che li precede, essa è talmente semplicistica che non val la pena di soffermarvisi troppo. Tali progressi *relativizzano* le concezioni e le teorie piú vecchie, ma non ne sopprimono il senso autentico.

Riserve. Il quadro tracciato nelle sue grandi linee è vero. La scienza è davvero una sola. Tuttavia questo problema è ricchissimo di sfumature. Malgrado la loro complessità non è possibile passarle del tutto sotto silenzio, con il rischio di arrivare a una visione troppo semplicistica dell'insieme del problema.

Una prima può essere quella che riguarda la categoria delle leggi della natura, che costituisce la disciplina detta termodinamica, cioè grosso modo la scienza delle relazioni tra calore ed energia. Queste leggi, va da sé, sono compatibili con le altre leggi della fisica. Ma nella forma che prendono in termodinamica non-statistica esse fanno ricorso, per descrivere la materia, a nozioni molto distanti da quelle che per lo stesso scopo utilizza tutto il resto della fisica. In larga misura la termodinamica statistica effettua in modo soddisfacente l'assimilazione reciproca di questi due insiemi di nozioni. La riduzione cosí ottenuta della termodinamica alla fisica *classica* è tuttavia lungi dall'essere totale. La termodinamica statistica comporta infatti elementi di soggettività che sono assenti dalla cosiddetta fisica classica. Indubbiamente si sono fatti in questi ultimi tempi grandi progressi nella ricerca di nuovi punti di vista in grado di eliminare tali elementi soggettivi. Ma è ancora troppo presto perché si possa contare sull'unanimità dei fisici a questo riguardo.

Una riserva analoga va fatta a proposito di un campo immenso della scienza contemporanea. Si tratta niente meno che di tutta la fisica *quantistica*: cioè della disciplina che fondamentalmente sta alla base del nostro sistema attuale di conoscenza scientifica. Già si è indicata la causa della necessità di una tale riserva. Ma per la sua importanza e complessità, l'argomento supera di gran lunga i limiti di questo paragrafo e costituirà il centro della discussione nel § 4.

Infine, un ultimo esempio di una concezione affascinante ma semplicistica consiste nel credere che una buona conoscenza *d'insieme* delle leggi fondamentali basti a spiegare *a grandi linee* i fenomeni osservati. Questa concezione è seducente, perché amiamo le sintesi eleganti. D'altra parte ha una sua parziale verità. Ma sarebbe un grave errore erigerla a verità prima e credere che i piccoli dettagli quantitativi dei fenomeni, quelli che si è soliti chiamare «accidenti», abbiano solo un'importanza del tutto secondaria per quel che riguarda la spiegazione dell'universo. Un celebre esempio è fornito dall'esistenza di un «livello eccitato» del carbonio 12 situato a 7,82 Mev. Tutti i nuclei atomici dispongono di molti «stati eccitati». Che il carbonio ne abbia uno a 7,82 Mev piuttosto che altrove è dunque un tipico «accidente», calcolabile (almeno in linea di principio), ma sprovvisto di particolare significato. Tuttavia è stato dimostrato che in assenza di questo livello nessun elemento pesante (piú pesante del berillio o del boro) si sarebbe formato nell'universo. In particolare non esisterebbero né carbonio né ossigeno... né per conseguenza alcun essere vivente, né quindi alcuna «persona», che per mezzo della propria riflessione cosciente possa conferire un significato alla nozione di universo.

Antimateria e cosmologia. Tali relazioni tra il piccolo e l'immenso conducono a considerare con particolare attenzione due recenti scoperte. In un tentativo di comprensione sintetica non si può forse attribuire loro la posizione piú centrale; non potrebbero affatto, in altri termini, essere la chiave di volta di alcun «sistema». Tuttavia esse fanno cadere dei tenaci pregiudizi, contribuendo cosí considerevolmente ad ampliare i nostri orizzonti. Si tratta da una parte della scoperta delle antiparticelle e dall'altra dei recenti progressi in cosmologia e cosmogenesi. Si potrà constatare, come già fatto sopra, che pur riguardando fenomeni e scale di dimensione cosí diverse, queste due questioni curiosamente non sono senza rapporti.

Nei grandi «acceleratori» come quelli del Cern (Ginevra) o del Fermilab (Illinois) si riescono a imprimere a delle particelle, ad esempio a nuclei d'idrogeno, chiamati anche «protoni», velocità molto vicine a quelle della luce; e questo in modo tale che i protoni abbiano un'energia molto piú grande della loro massa a riposo (misurata con appropriate unità). Che cosa si può osservare quando questi protoni entrano in collisione ad esempio con altri protoni? In genere si vedono emergere dall'interazione non soltanto le due particelle iniziali, ma anche molti altri oggetti, che sono ugualmente delle «particelle», che possiedono una massa a riposo, ecc. Tutti questi oggetti sono stati *creati* a spese dell'energia cinetica del protone incidente, secondo l'identità energia-massa scoperta da Einstein. Qualche volta succede pure che tra questi oggetti creati si possano osservare uno o piú protoni. In questo caso si osservano però anche un certo numero di particelle con una massa a riposo uguale a una carica elettrica opposta a quelle che caratterizzano il protone. Questi ultimi oggetti si chiamano «antiprotoni». Nello stesso modo a ciascun tipo di particelle corrisponde un altro tipo di particelle, denominate «antiparticelle» delle prime (certe particelle neutre sono antiparticelle di se stesse).

In una regione dello spazio dove non ci sono sistemi che restino a lungo senza contatto con gli altri, non si possono trovare allo stato naturale contemporaneamente delle particelle e le loro antiparticelle, a meno che si tratti di

un sistema in stato esplosivo. Questo è dovuto al fatto che quando una particella e un'antiparticella di tipo simile entrano in collisione, esse si annichilano reciprocamente, fornendo energia (sia cinetica, sia formata da energia di massa di particelle piú leggere). Cosí nella nostra porzione di universo le particelle protoniche sono positive (protoni) e le particelle elettroniche sono negative (elettroni). Antiprotoni (negativi) e antielettroni (positivi) non possono essere ottenuti che artificialmente (con una sola eccezione: le antiparticelle create dalla radiazione cosmica bombardando gli alti strati dell'atmosfera).

Una breve digressione: risulta utile classificare le interazioni fondamentali in quattro tipi. Per ordine di forza crescente si tratta di interazioni gravitazionali, interazioni deboli (responsabili di talune radioattività), interazioni elettromagnetiche, conosciute da tutti, e interazioni forti, il cui raggio è molto corto (10^{-13} cm), ma la cui intensità è assai notevole, assicurando in tal modo la coesione dei nuclei. Ad eccezione delle interazioni deboli, tutte queste forze sembrano perfettamente simmetriche in rapporto alla conversione virtuale simultanea di tutte le particelle in antiparticelle. Ne risulta che un mondo fatto di antimateria sarebbe strutturalmente molto simile al nostro, e che taluni oggetti celesti sono forse costituiti di antimateria.

La scoperta delle antiparticelle ha arrecato un colpo decisivo alla vecchia concezione atomistica, secondo cui esisterebbero delle particelle elementari indistruttibili ed eterne, di cui i comuni oggetti sarebbero dei semplici aggregati. Val la pena di notare che questa scoperta di grande portata era stata prevista dalla fisica teorica. Molto prima che fossero osservate sperimentalmente, l'esistenza delle antiparticelle era apparsa ai teorici come forse l'unico modo concepibile di armonizzare tra di loro le prospettive delle relatività e della meccanica quantistica: ciò per ragioni di carattere essenzialmente matematico, che non possono essere qui illustrate.

Come è noto, le concezioni relativistiche svolgono anche un ruolo di primo piano nella nostra comprensione della struttura dell'universo. Si tratta in questo caso soprattutto della relatività generale, teoria anch'essa notevolmente matematizzata, che permette di classificare in un numero ristretto di tipi i «modelli cosmologici» accettabili a priori. La scelta tra questi modelli resta naturalmente questione di osservazione.

L'osservazione in questione ha messo in evidenza dapprima (Hubble) il fatto fondamentale di uno spostamento verso il rosso delle righe spettrali delle galassie, che aumenta con la distanza. Ancora una volta, questo spostamento è generalmente interpretato come rivelatore di una dilatazione dell'universo. Per renderne conto bisognava fare l'ipotesi o di una «creazione continua» o di un «inizio». Il fisico teorico Gamow ha proposto una teoria dell'evoluzione fondata sulla seconda concezione (e, naturalmente, anche sull'ipotesi che le leggi non sono fondamentalmente mutate). Da questa teoria ha dedotto l'esistenza di fenomeni, alcuni dei quali, osservati (radiazione intergalattica in equilibrio termico a 4 gradi Kelvin) risultano difficilmente spiegabili da qualsiasi altra teoria. Cosí l'ipotesi di un «inizio» (se sia preceduto o no dall'«esistenza» di un universo in contrazione, o da un qualche essere

diverso, è una questione che sfugge alla scienza) è attualmente considerata verosimile.

Negli ultimi vent'anni sono stati scoperti nell'universo oggetti notevoli, come i quasar e le pulsar. La fisica delle particelle elementari rende conto senza ambiguità delle seconde («stelle di neutroni») e offre possibilità d'interpretazione dei primi. La relatività generale permette anche di predire l'esistenza di strutture notevoli come i «buchi neri», residui di antiche grandi stelle, nei quali tutto ciò che penetra deve consumarsi senza possibilità di ritorno: è possibile che certi fenomeni celesti recentemente osservati non siano privi di rapporti con l'esistenza di tali oggetti.

La fisica contemporanea è cosí divenuta lo strumento di analisi teorica privilegiato non soltanto dai chimici, ma anche dagli astronomi, realizzando in tal modo l'unità di scienze tanto diverse in relazione al loro oggetto.

Resta, è vero, la biologia, i cui fenomeni sono tanto complessi che non deve stupire la particolare difficoltà che può esserci nel ricollegarli, attraverso connessioni deduttive, ai fenomeni elementari. Cosí si rende necessario correggere qui un errore banale, ma commesso purtroppo troppo sovente. Siccome i recenti progressi della biologia attirano a giusto titolo l'attenzione, certi personaggi, diciamo dal giudizio «rapido», credono che la fisica sia *scavalcata* dalla biologia, cioè che certi concetti, considerati fondamentali in fisica, non siano piú validi in biologia. Beninteso, ciò che mostrano la biologia molecolare e la teoria dei sistemi dissipativi è esattamente il contrario: secondo ogni evidenza non c'è legge della biologia che trasgredisca quelle della fisica; si è autorizzati a pensare che le leggi e i concetti della fisica siano universali, e comunque, per quel che riguarda i corpi viventi sembrano interamente validi. Di vero sembra che ci sia soltanto il fatto, d'altra parte interno alla fisica, che l'irreversibilità macroscopica non è deducibile dalla reversibilità microscopica senza l'introduzione di elementi esterni.

Se quindi gli esseri viventi sono tanto organizzati, è in virtú di una evoluzione che – per quanto improbabile abbia potuto essere – ha sempre rispettato le leggi della fisica. Analogamente le forze che assicurano la coesione e il moto delle molecole e degli atomi sono forze puramente fisiche. Ci si può d'altra parte domandare perché le forze fisiche, che sono semplici e che in altri casi (sistema solare, interno delle stelle, ecc.) producono solo aggregati poco strutturati, diano origine a strutture tanto complesse e ordinate. La risposta a questo problema è sicuramente difficile e non ancora pienamente conosciuta. Se ne può tuttavia fin d'ora notare un elemento indubbiamente valido. Mentre le condizioni locali, temperatura, pressioni, ecc. non permettono in generale agli effetti quantistici di manifestarsi molto nel cosmo, sulla Terra queste condizioni sono per caso tutte favorevoli alla loro manifestazione. Gli effetti quantistici, come indica il loro nome, fanno intervenire dei numeri interi. In altri termini, mentre tra la figura che descrive la Terra intorno al Sole e quella che descrive la cometa di Halley è possibile concepire quante figure intermedie si vogliano – tutte realizzabili in linea di principio –, niente di simile è possibile per quel che riguarda gli atomi e le molecole. Non c'è ne-

cessariamente una differenza qualitativa tra uno stato quantico di una molecola e un altro. Questo conferisce a ciascun stato quantico una *individualità* che nessun stato classico può avere; nonché una stabilità (relativa peraltro, ma reale) che uno stato classico non ha, salvo condizioni particolari. Si capisce dunque molto bene – e il calcolo lo conferma – che possano formarsi associazioni di molecole in stati individualizzati e quasi-stabili che abbiano essi stessi una netta individualità, e si può dunque immaginare che esse – eventualmente – si riproducano con una buona stabilità (questo fattore di stabilità, lo si noti qui a titolo informativo, non esclude pertanto numerosi errori che la selezione naturale da parte sua corregge).

Beninteso, si può legittimamente parlare il linguaggio della cosiddetta fisica classica quando si dà per scontata l'individualità delle molecole costituenti e quando non si cerca, d'altra parte, di calcolare esattamente le forze che servono a tenerle legate; quando, in altri termini, ci si limita a uno studio semiqualitativo (che costituisce una tappa indispensabile in problemi tanto complessi). È quanto con piena ragione fanno, in generale, i biologi molecolari della nostra epoca. Nessuno di essi si sogna di negare che le leggi alle quali obbediscono le molecole da essi studiate siano le leggi della fisica, né, di conseguenza, che tali leggi siano fondamentalmente quantistiche; né infine negano il ruolo di fatto, anche se molto indiretto, svolto da questo carattere quantistico della fisica di base per quello che riguarda l'individualità, cioè per quel che riguarda la stabilità strutturale dei corpi viventi.

3.2. Ruolo della matematica nella fisica.

La matematica forma incontestabilmente l'ossatura delle scienze esatte. È interessante studiare, anche sommariamente, questa funzione.

Nel tempo essa si è evoluta. Se si prende in esame l'«era classica», circa fino al 1900, vi si constata l'esistenza di una fisica abbastanza concreta. La dominano in effetti due concetti, materia e campo, nessuno dei quali richiede per essere compreso il ricorso alla matematica (in particolare, Faraday introdusse, come è noto, la nozione di campo attraverso il concetto di «linee di forza», che s'impone all'intelletto con molta chiarezza e senza astrazioni). Non per questo le matematiche dell'epoca non hanno un ruolo di primo piano nella fisica. Questo ruolo è però *secondario* nel senso etimologico del termine: *viene dopo* la definizione del concetto e permette di servirsene. La matematica in particolare consente di rendere quantitative teorie che certamente potrebbero essere *descritte* senza il suo aiuto, ma dalle quali senza di essa sarebbe impossibile trarre delle conclusioni quantitative. Di esse, di conseguenza, non sarebbe possibile in generale verificare l'esattezza. Con queste funzioni, relativamente modeste, la matematica già esercita tuttavia un notevole potere sulla fisica attraverso le nozioni di semplicità e universalità. Cosí ad esempio, l'immediato successo presso i dotti dell'epoca dell'ipotesi copernicana non è attribuibile unicamente a un migliore accordo con l'esperienza (la teoria geocentrica di Tolomeo da questo punto di vista non andava male): ma va al-

meno in parte ascritta alla sua «eleganza matematica». Analogamente, se il potere di convinzione della teoria della gravità di Newton trionfa cosí rapidamente – benché l'idea di un'azione a distanza sia oscura – è, evidentemente, grazie all'universalità della legge proposta. Già in quest'epoca la matematica permette di cogliere la nozione di *struttura* indipendente dai supporti fisici: permette ad esempio di riconoscere le stesse fondamentali strutture in tutti i *fenomeni ondulatori*. Si noti a tal proposito che è di per sé notevole che le strutture fondamentali siano matematicamente intelligibili. A chi domandasse «Qual è la piú grande scoperta della fisica?» si risponderebbe forse «Il fatto che la struttura dei fenomeni è parallela a quella del nostro spirito».

Infine – e non è certo la cosa meno importante – c'è un'altra ragione del fatto che la matematica è stata fin dall'inizio un elemento unificante della fisica: la matematica non costituisce soltanto un linguaggio comodo, in realtà rappresenta un linguaggio *piú* un ragionamento. Si è spesso sottolineato che è soprattutto in questo che la matematica è insostituibile. Cosí, ad esempio, a un uomo che non capisce la matematica si può benissimo *descrivere*, nel linguaggio quotidiano, la legge classica della rifrazione della luce (legge di Snell-Descartes). Si traccia ad esempio una sfera di raggio unitario con centro il punto d'impatto del raggio luminoso. Dai due punti d'intersezione di questa sfera con il raggio luminoso si abbassano le perpendicolari sulla normale al piano di rifrazione. Si enuncia allora la legge affermando che questi due segmenti sono coplanari e che il loro rapporto è uguale a una costante, indipendente dall'angolo d'incidenza. Alla stessa persona può essere descritto in modo analogo il principio di Fermat senza far ricorso alla matematica. Basta dire che la velocità della luce dipende dal mezzo nel quale si propaga e che il cammino del raggio luminoso per andare da un punto all'altro è sempre quello che esige il tempo minimo (o massimo). Tuttavia, senza matematica è impossibile spiegare a questa persona che i due enunciati precedenti, in apparenza cosí diversi, di fatto dicono precisamente la stessa cosa e che per il problema sono rigorosamente equivalenti. Cosí l'ignoranza della matematica impedisce di passare da una descrizione, istruttiva in questo o quel caso, a una diversa descrizione che in certi altri casi illumina meglio. Tutti i fenomeni naturali sono in realtà fortemente interconnessi: la nostra impossibilità di mutare il punto di vista si traduce in pratica in una incapacità a cogliere i legami profondi che uniscono fenomeni tanto diversi della natura e costituisce dunque un ostacolo a ogni autentica comprensione.

Va aggiunto però, per sfumare quest'osservazione, che l'eccesso opposto è quasi altrettanto paralizzante. A voler conoscere soltanto le equazioni, a voler ignorare i concetti che ne sono alla base e i problemi che essi sollevano, ci si precluderebbe in modo altrettanto radicale l'autentica comprensione. Ci si precluderebbero soprattutto molte possibilità di porre dei problemi, alcuni dei quali sono interessanti e possono condurre a progressi.

Questo resta vero anche oggi. Tuttavia se si volge l'attenzione all'era moderna – che, per la scienza, è il nostro secolo – vi si constata che, oltre alle funzioni già descritte, la matematica ne assume un'altra, la cui importanza

va crescendo. Sempre piú la matematica *serve a definire gli stessi concetti che si cerca di porre in corrispondenza con il fenomeno osservato.*

Si può, volendo, far risalire la «preistoria» di questo metodo a un passato ormai lontano. La formazione del concetto di energia ne è in un certo senso una delle prime applicazioni. L'attingere i propri concetti base – come vettore di stato, algebra di osservabili, ecc. – direttamente dalla matematica stessa costituisce tuttavia un aspetto realmente originale della fisica odierna. Per di piú, essa li attinge di preferenza dalla matematica moderna, dove le nozioni corrispondenti sono definite in modo astratto, cioè, ad esempio, come insiemi muniti di strutture definite o come elementi di tali insiemi. Questo fatto potrebbe indurre ad aderire alla lettera alle concezioni di Pitagora, il quale affermava che l'essenza stessa delle cose è costituita dai «numeri» (lo si può tradurre con «matematica»). Ciò, è ovvio, non fa che rendere ancora piú aleatori quei tentativi di descrizione dettagliata dei fondamenti della fisica moderna, che vorrebbero fare a meno di questa stessa matematica. Una tale verità può sembrare scoraggiante a chi non è un matematico. Un apprezzabile conforto può essere tratto dal fatto che la matematica necessaria per la comprensione delle idee-forza della fisica è molto semplice. Inoltre, e non bisogna stancarsi di ripeterlo, se la matematica è indispensabile, non è tuttavia sufficiente (contrariamente, questa volta, a quel che pensava Pitagora). Poca matematica, ma ben scelta, e una buona dose di riflessione, ecco ciò che è necessario, per penetrare a una qualche profondità nella problematica del nostro tempo.

Infine se tutto quello che qui è stato detto sul ruolo della matematica, si applica perfettamente alla fisica *in generale*, bisogna aggiungere un'osservazione molto importante riguardo alla fisica quantistica, nella quale questo ruolo è particolare. Si è già visto, e piú avanti vi si tornerà, che la meccanica quantistica (almeno quella che oggi s'insegna e utilizza) non può pretendere di esporre *ciò che è*, ma cerca soltanto di descrivere *ciò che si osserva*. La matematica che essa utilizza non può dunque essere considerata – nemmeno implicitamente – di portata ontologica. Al massimo si può legittimamente ipotizzare l'esistenza di *corrispondenze* (ma di quale natura resta un mistero) tra le sue strutture e quelle della realtà indipendente. Questa osservazione verrà sviluppata nelle prossime pagine.

4. *La fisica e i rapporti dell'uomo con la natura.*

4.1. Democrito e Pitagora.

Oggi dunque, anche grazie in primo luogo alla fecondità dei principî di base della meccanica quantistica, il potere di sintesi della fisica sembra, almeno di diritto, essere totale o quasi.

In tali condizioni è evidentemente essenziale studiare la natura della conoscenza fisica in generale e – in particolare – quella di tali notevoli «principî quantistici». Non li si formula qui con particolari piú numerosi di quanto già fatto: il compito risulterebbe troppo tecnico. È importante enuclearne almeno i concetti di base.

All'inizio di un'impresa cosí generale sembrano a priori aperte due strade ben simboleggiate da due nomi. Due nomi che risultano emblematici, poiché ciascuno evoca un atteggiamento di pensiero, che fu fecondissimo nella fisica: Democrito e Pitagora. Il nome di Democrito è intimamente legato alla dottrina dell'atomismo, secondo la quale l'Essere si trova disseminato in una moltitudine quasi infinita di frammenti semplicissimi, ben localizzati nello spazio, raggruppati in un piccolo numero di tipi, e in grado di esercitare gli uni sugli altri soltanto delle interazioni ugualmente localizzate, o di contatto, o almeno decrescenti con l'aumentare della distanza. Influenze che del resto sono in parte forze attrattive, che hanno perciò tendenza a costruire con tali «atomi» degli aggregati, essi stessi approssimativamente localizzati. Secondo questa dottrina, tali aggregati costituirebbero tutti gli oggetti che esistono nell'universo. Quanto a Pitagora, il suo nome evoca la dottrina, ugualmente degna di rispetto, ma del tutto diversa e anche opposta, secondo la quale «i numeri sono l'essenza delle cose». Senza dubbio questo stesso nome evoca anche una certa mistica dei numeri, che la fisica ha a ragione abbandonato da lungo tempo. Si può tuttavia convenire di dimenticare la componente mistica del pitagorismo e attribuire il nome di atteggiamento pitagorico alla concezione, che è insieme metodo e dottrina, che vede nella matematica l'essenza ultima del reale. È quanto si farà qui.

Con tali definizioni terminologiche è possibile ora studiare il problema seguente. È possibile interpretare in termini di atomismo democriteo la fisica attuale, quella a cui i fatti sperimentali studiati e analizzati nel corso dei secoli hanno oggi portato, la fisica quale insomma si scopre ai nostri occhi (senza che essa in nulla sia una scelta arbitraria dello spirito)? O è invece interpretabile secondo l'atteggiamento pitagorico? Oppure, si basa su una sottile mescolanza dei due? O infine si tratta di tutt'altra cosa, non riducibile ad alcuna di queste due concezioni di base? Beninteso, queste domande hanno senso soltanto se atomismo e pitagorismo sono intesi non come ricette di calcolo, ma come tentativi di descrizione della realtà indipendente. È questo il punto di vista qui adottato.

Detto questo, c'è nella fisica attuale un po' di atomismo democriteo? In verità, sarebbe facile crederlo, a giudicare da quel che si scrive nelle scienze e sulla scienza. Il riferimento ultimo nelle piú diverse discipline non pare forse essere l'atomo, o meglio la particella elementare? L'astrofisico non parla forse di gas di particelle ionizzate? Il biologo molecolare non descrive le macromolecole di cui si occupa come edifici complessi di atomi, che hanno gli uni in rapporto agli altri delle situazioni ben precise?

Il linguaggio dell'atomismo è sicuramente di prodigiosa fecondità. Ha inoltre l'immenso vantaggio di venire in aiuto all'immaginazione scientifica con la sua impareggiabile limpidezza (che cosa è piú semplice da visualizzare di un «piccolo granello»?) E tuttavia, malgrado tutte queste apparenze, a dispet-

to anche dell'opinione contraria della gran maggioranza degli scienziati che lavorano in campi lontani da quello della microfisica, l'atomismo democriteo non è scientificamente la verità di ciò che *è*.

Questo atomismo è solo un modello. E dire questo non significa ancora enunciare solo una restrizione sufficientemente forte, perché ci sono modelli di molti tipi. Alcuni possono essere considerati buone approssimazioni del reale: è il caso del sistema isolato nella meccanica classica. Nessun sistema è rigorosamente isolato, ma si constata che in molti casi, in cui le interazioni di un sistema fisico con il mondo esterno sono deboli, il fisico non commette un errore grave ponendole uguali a zero. Egli costruisce cosí un modello vicino al sistema. In questo tipo di modello si fa l'ipotesi che, almeno sul piano della qualità, i concetti di base del modello riflettano adeguatamente la realtà cosí come essa è veramente, indipendentemente dalla ragione umana. Soltanto sul piano quantitativo si suppone che questo tipo di modello si allontani in qualche misura dalla realtà, che è complessa. Ma esistono anche modelli di un secondo tipo, modelli che nella pratica risultano certamente utili, ma a proposito dei quali è possibile dimostrare che i concetti base da essi utilizzati *non* riflettono la realtà. Un semplice esempio è fornito dal modello della sfera celeste e dell'eclittica nel quale si dice che il Sole percorre un certo cammino sulla sfera delle stelle fisse. Ora, la fisica contemporanea ha insegnato che l'atomismo democriteo esteso agli elettroni è un modello di questo secondo tipo. Non può trattarsi di un modello del primo tipo.

Il fatto ora sottolineato, cioè l'impossibilità di considerare l'atomismo democriteo se non come un modello del secondo tipo, i cui concetti di base *non* riflettono la realtà come è in sé, è un fatto noto a chiunque sia iniziato alla teoria, anche elementare, dell'atomo, cioè alla meccanica quantistica. È dunque inutile insistere sulle prove a suo favore, se non per rispondere a un argomento portato talvolta per metterlo in questione. Tale argomento si fonda sulla circostanza che l'impossibilità segnalata si riferisce a una teoria: ora se si ha fiducia in questa teoria è soltanto perché è verificata dall'esperienza o, in altri termini, perché l'esperienza verifica le predizioni da essa fornite circa i risultati delle diverse misure che si possono fare. Ma si può pensare che anche altre teorie, fondate su un diverso sistema di concetti di base, forniscano queste stesse predizioni. Schematicamente, quest'argomento consiste nel far notare che fra tali altre teorie – ammesso che esistano – qualcuna potrebbe forse descrivere la realtà in modo compatibile con l'atomismo democriteo.

La risposta a questo argomento dev'essere data in due tempi. In primo luogo: è pur vero che esistono delle teorie fondate su concetti molto diversi da quelli che servono di base alla meccanica quantistica (che è la teoria correntemente in uso) e che forniscono tuttavia le stesse predizioni sperimentali (almeno nei campi in cui queste teorie sono in grado di farlo). Si tratta di teorie molto complesse, conosciute con il nome tecnico (di cui sarebbe inutile e fastidioso esporre qui l'origine) di «teorie a variabili nascoste». In secondo luogo: malgrado l'esistenza di queste teorie l'atomismo democriteo non

è conciliabile con la fisica contemporanea, perché nessuna teoria a variabili nascoste è compatibile con le concezioni di base di questo atomismo appena ricordato. Il punto d'incompatibilità è quello che riguarda le influenze a distanza. Si è potuto infatti dimostrare con completa generalità che qualsiasi teoria a variabili nascoste, nella quale le influenze dirette fra le particelle diminuiscono con il crescere della distanza (come è necessario che facciano, perché l'atomismo abbia un senso), implica l'esistenza di alcune disuguaglianze tra quantità accessibili all'esperienza (disuguaglianze di Bell). Ora, le predizioni della meccanica quantistica violano queste disuguaglianze. In effetti le sole teorie a variabili nascoste che riproducano esattamente le predizioni della meccanica in questione sono quelle dette «non-locali». Queste implicano l'esistenza d'influenze reciproche molto forti tra particelle che abbiano interagito nel passato, con la conseguenza che queste particelle, benché la teoria in discussione si sforzi di considerarle distinte e separate, persistono di fatto nel formare un tutto: e questo anche nei casi in cui la teoria le descrive come lontanissime l'una dall'altra e l'esperienza conferma l'estensione spaziale del «tutto» che esse insieme costituiscono. Come si vede, si tratta di condizioni che rendono privo di senso il linguaggio dell'atomismo *filosofico*.

Cosí, dunque, la prima delle due vie che a priori era opportuno esplorare, quella dell'atomismo democriteo, si rivela a posteriori incompatibile con i fatti. Resta dunque da esplorare l'altra, quella evocata dal nome di Pitagora. I numeri – cioè la matematica – sono l'essenza di tutte le cose? Una tale questione è, beninteso, molto vasta e il suo studio sistematico porterebbe troppo lontano. È dunque legittimo restringerla grazie ad alcune considerazioni generali. Si può forse in questo modo eliminare oggi un'idea che nel passato ha sedotto molti pensatori. Si tratta di coloro che, matematici o filosofi, di fronte all'incertezza che intacca ogni conoscenza proveniente dai sensi, ritennero che le sole verità eterne e oltrepassanti l'uomo fossero le verità matematiche. Certo, una tesi cosí generale non può a rigore essere confutata. Al momento attuale, però, tutto un insieme di osservazioni – provenienti sia dagli studi di psicologia infantile, sia dalle nostre conoscenze biologiche o genetiche – convergono e tendono a dimostrare che la matematica riflette le capacità operative dell'essere umano, si tratti del bambino o, secondo altri, dell'uomo primitivo.

Cosí ai nostri giorni i difensori di quello che qui sopra è stato chiamato «l'atteggiamento di Pitagora» guardano piú dalla parte della fisica matematica che da quella della matematica pura. La loro tesi suonerebbe all'incirca cosí: «Certo esiste una realtà indipendente dall'uomo. Questa realtà in linea di principio è conoscibile con i mezzi della fisica, ma trova adeguata espressione non nei concetti della vita quotidiana, bensí nelle equazioni della fisica matematica». Come si è visto, era questa la tesi di Einstein. È esatta?

Iniziando un esame del problema va subito riconosciuto che lo studio di campi essenziali e fondamentali della fisica fornisce argomenti molto persuasivi in favore di questa tesi. Il ruolo della matematica nella fisica non è affatto limitato a quello di una scrittura stenografica: ci sono dei concetti come

quelli di spazio curvo, o di «vettore di stato», che possono essere esposti e capiti in modo adeguato soltanto facendo ricorso alla matematica. Piú in generale, giocano a favore della tesi in questione tutto l'elettromagnetismo classico, tutta la relatività generale e quindi buona parte della cosmologia. Bisogna aggiungere che le carenze dell'atomismo filosofico forniscono, per contrasto, argomenti in suo favore. Cosí, per esempio, non è possibile asserire con completa generalità di una qualsiasi particella che è un «atomo», che è cioè indistruttibile e indivisibile (requisito essenziale per poter costituire «l'elemento ultimo della realtà»): lo scontro di due particelle può infatti produrre altre particelle, o piú esattamente, particelle e antiparticelle associate. Come si può dunque parlare ancora di una «conservazione della materia»? Risposta: o identificando quest'ultima con l'energia, o assimilandola a una differenza tra un numero di particelle e un numero di antiparticelle. In entrambi i casi si è portati a conferire il primato a una quantità *astratta*, di natura essenzialmente matematica.

Non sorprende quindi constatare che la grande maggioranza dei fisici teorici contemporanei aderiscano – implicitamente o meno – a quello che è stato qui sopra chiamato «l'atteggiamento di Pitagora» e non siano lontani dal far propria la frase di Jeans «Dio [s'intenda: la Realtà ultima] è un matematico» (s'intenda: la sua essenza è la fisica matematica).

Questo atteggiamento filosofico risulta, ancora una volta, estremamente comprensibile di fronte a certi dati della fisica piú recente. Si scontra tuttavia con un'obiezione abbastanza evidente, alla quale si è accennato alla fine del § 3 e che bisogna ora discutere. Si tratta del fatto che tutte le teorie moderne di fisica matematica – la teoria della matrice S, la teoria quantistica dei campi, o qualsiasi altra costruzione con gli stessi oggetti – definiscono lo stato di un sistema mediante *il suo modo di preparazione*. Una tale definizione è considerata dagli esperti come di gran lunga la piú soddisfacente. E bisogna anche dire come *l'unica* soddisfacente. Tutto questo a buon diritto. Infatti, se appena si approfondiscono i dettagli tecnici, si constata che si tratta della ricetta infallibile per eliminare dei paradossi che altrimenti rischiano di sorgere.

Quest'ultimo punto merita un commento. Per ragioni filosofiche, si può senza dubbio cercare di eliminare dalla teoria il carattere antropomorfo della nozione di «preparazione di uno stato». Si potrebbe ad esempio essere tentati di dire che tutti gli insiemi statistici di sistemi sono regolati da principî quantistici fondamentali, siano questi insiemi conosciuti o soltanto definiti da proprietà ben specificate, ma che si ammette che nessuno conosca. In particolare, sarebbe questo il caso di un insieme di sistemi, ciascuno dei quali incorpori uno strumento di misura, considerato come un semplice sistema fisico finito: le posizioni finali delle lancette degli strumenti avrebbero valori intrinseci e ciascuno dei valori possibili definirebbe un sottoinsieme, indipendentemente da qualsiasi intervento di un qualunque osservatore. Una tale concezione avrebbe l'evidente vantaggio di unire la concezione di una realtà indipendente dall'uomo (e in linea di principio descrivibile dalla fisica) con l'idea, giustificata dall'esperienza, che i principî quantistici hanno validità affatto generale. Sfortunatamente la sua adozione conduce direttamente a incoerenze tecniche che si possono riscontrare nelle opere specialistiche. Il programma consistente nel tentare di eliminare queste incoerenze mediante l'attribuzione agli strumenti di proprietà idealizzate che riflettono la loro complessità, si è rivelato dei piú ardui e d'incerto successo. I fisici teorici hanno dunque completamente ragione quando insistono a definire gli stati dei sistemi mediante il loro modo di preparazione riferito, s'intende, al libero arbitrio del fisico sperimentatore, e quando rifiutano d'interessarsi al meccanismo complesso di questa stessa preparazione.

Ma gli stessi fisici avrebbero meno ragione, se, pensando alla realtà ultima e considerandola indipendente dall'uomo, credessero di poter affermare che le equazioni della fisica matematica sono in grado di descriverla, o almeno di darne una rappresentazione asintoticamente tendente a una rappresentazione fedele, man mano che aumentano le nostre conoscenze. Pensando questo, essi dimenticherebbero il ruolo eminente svolto dalle nozioni di preparazione degli stati e di osservazione dei sistemi negli stessi enunciati dei principî di base della loro teoria. Va sottolineato che questo ruolo essenziale delle nozioni di preparazione e di osservazione è una novità introdotta all'inizio del secolo XX dalla «rivoluzione quantistica». I principî di base della fisica classica erano formulati senza riferimento a queste nozioni. È vero che era sempre possibile formularli mettendo in gioco anche tali riferimenti; ma quel che qui importa è che tali riferimenti non erano nei loro confronti essenziali e nemmeno praticamente utili (salvo che un po' nella relatività). Nella fisica quantistica esse sono indispensabili nella pratica e si può anche sostenere l'idea che siano essenziali in linea di principio. Le circostanze, qui sommariamente descritte, mostrano che la seconda via che si presentava inizialmente, quella dell'«atteggiamento di Pitagora», fondata ancora una volta sull'idea di una realtà *indipendente* conoscibile e matematicamente descrivibile, non è nemmeno essa una via regia per la conoscenza del mondo. È innegabile che presenti una parte di verità. Ma non costituisce da sola una visione globale della verità, poiché non viene eliminato dal formalismo l'esplicito riferimento all'uomo o a strumenti *che considera come tali*; e poiché, ancora una volta, la matematica non funziona qui da referente esclusivamente ontologico.

4.2. Oggettività forte e debole.

Bisogna dunque cercare una terza via. È chiaro che in questa ricerca si terrà conto delle difficoltà incontrate nell'esplorazione delle prime due. In particolare pare impensabile non tener conto del fatto pratico appena descritto: cioè che i principî di base della fisica atomica e delle particelle fanno esplicito riferimento a nozioni di preparazione e osservazione, in ultima analisi, all'attività e alla sensibilità umane. Questo conduce niente meno che a una ridefinizione dell'oggettività.

Bisogna insistere su questo punto: è fondamentale, elementare, ma soven-

te trascurato. Siccome la nozione di oggettività è la nozione di base di tutta la scienza, le contraddizioni che la riguardano hanno le piú gravi conseguenze. Sono in grado di snaturare completamente il messaggio d'ordine generale di cui è portatore lo sforzo della ricerca scientifica.

Per formulare adeguatamente le necessarie distinzioni, la cosa piú semplice è quella di adottare come punto di partenza il fatto che gli scienziati stessi non attribuiscano tutti lo stesso senso alla parola 'scientifico'. Si tratta di un'eredità dello sviluppo storico delle idee riassunte nel § 3. Nei ricercatori si trovano infatti partigiani di entrambe le concezioni principali. Secondo la prima, quella del realismo fisico, esiste una realtà fisica indipendente, che può essere approssimata dalla nostra scienza. Lo sperimentatore umano non è che un sistema fisico tra gli altri; è anzi un sistema fisico estremamente particolare, che esiste solo in un'infima parte dell'universo. Sarebbe dunque assurdo fondare la fisica sulle proprietà di un tale sistema, e ancora piú assurdo – se possibile – fondarla sulle proprietà piú sottili di questo sistema (libero arbitrio, coscienza). Gli scienziati che aderiscono al realismo fisico esigono dunque, del tutto legittimamente, che l'enunciato di un principio fisico fondamentale non faccia nessun riferimento, nemmeno implicito, alle capacità operative e di osservazione dell'essere umano. Tale è, e tale dev'essere, la loro concezione dell'oggettività. Si chiamerà *oggettività forte* questa concezione.

Gli scienziati che adottano la concezione opposta sottolineano il fatto che ci si deve sempre accuratamente guardare dal costruire delle frasi prive di senso. Fanno notare che nei campi che si allontanano da quello della vita quotidiana un pericolo siffatto è sempre in agguato. A loro avviso l'unico modo sicuro di evitarlo è quello di definire in modo esclusivamente operativo le nozioni sulle quali si basano tali frasi. Siccome per questi scienziati la definizione di un concetto si riferisce esplicitamente a un'osservazione o a un atto umano, è chiaro che non possono definire l'oggettività di un'affermazione o di una nozione, in modo che risulti radicalmente indipendente dalle possibilità (o dalle impossibilità) operative della comunità degli esseri umani: che non possono insomma riferirla a una qualche nozione di «realtà» a sua volta non suscettibile di alcuna definizione operativa. Fondano di conseguenza la loro definizione dell'oggettività sulla comunicabilità. L'obiettività di un'affermazione consiste soltanto nella sua validità per un essere *umano qualsiasi*. Si chiamerà questa concezione dell'oggettività, *oggettività debole*. La differenza tra oggettività forte e debole sta nel fatto che un enunciato scientifico può essere debolmente oggettivo, pur riferendosi esplicitamente al libero arbitrio o alla sensibilità dell'essere umano in generale, a condizione che non venga fatto nessun riferimento alle facoltà e alle impressioni di un essere umano particolare.

La fisica contemporanea di base – con la sola eccezione, se si vuole, della relatività generale – è tutta costruita sulla nozione di oggettività debole. Certo, alcuni fisici pensano che dovrebbe essere possibile ricostruirla in termini di oggettività forte, e hanno anche fatto dei tentativi in questo senso. Sono giunti ad alcune interessanti conclusioni: si tratta ancora tuttavia di risultati molto parziali, che lasciano sussistere vaste zone d'ombra e anche alcune obiezioni di principio.

5. *Conclusione. Fisica e visione del reale.*

Molti fisici vorrebbero mantenere la fisica libera da ogni filosofia e pensano che ciò sia possibile. Le considerazioni precedenti mostrano che una tale speranza è vana, a meno che non si pensi che adottare un punto di vista deliberatamente operazionista non sia fare della filosofia: adottare cioè un punto di vista che scarta come priva di senso ogni nozione di realtà concepita come indipendente dalla collettività degli esseri umani. Un'indicazione in favore dell'idea che un tal punto di vista operazionista è in sostanza filosofico – e non puramente scientifico – è fornita dal semplice fatto che gli scienziati si sono schierati al suo riguardo su posizioni opposte, come piú sopra si è detto. A questo proposito è interessante qualche citazione. A favore del punto di vista operazionista si possono citare molti passi di Bohr, del tipo «La meccanica quantistica, il cui solo scopo è la corretta descrizione delle osservazioni», o ancora la seguente osservazione di Schrödinger: «Non esistendo per nessuno e, di conseguenza, inesistente». Sul lato opposto l'avversario piú noto e costante di tale punto di vista è stato – è facile indovinarlo – Albert Einstein, che, poco prima della sua morte, non esitava a scrivere: «C'è qualcosa come "lo stato reale" di un sistema fisico, che esiste obiettivamente, indipendentemente da ogni osservazione o misurazione e che in linea di principio si descrive con i mezzi di espressione della fisica (quali mezzi di espressione e, di conseguenza, quali concetti fondamentali si debbono utilizzare al riguardo... è cosa che secondo me ancora non si conosce). Questa tesi sulla realtà non ha il senso di un enunciato chiaro di per sé, a causa del suo carattere "metafisico". Ha soltanto il carattere di un programma. Tutti gli uomini si attengono tuttavia saldamente a questa tesi sulla realtà, compresi quei teorici quantistici che non pongono in discussione i fondamenti della teoria quantistica. Nessuno ad esempio dubita che il centro di gravità della Luna non occupi a un dato istante una posizione determinata, anche in assenza di un qualsiasi osservatore, reale o potenziale. È vero che questa tesi è arbitraria da un punto di vista strettamente logico, ma diventa ben difficile sfuggire al solipsismo se la si lascia cadere» [1953, p. 6].

Con queste affermazioni Einstein si colloca inequivocabilmente nel campo dei partigiani del realismo, e del realismo fisico. Altrove si è compromesso anche di piú e in modo piú preciso. Con Podolsky e Rosen ha infatti insistito sulla nozione di elemento di realtà e, implicitamente, ha associato questa nozione a quella di *località*. Per lui non solo la realtà era in linea di principio descrivibile con i mezzi della fisica, ma inoltre la legge (relativistica) della propagazione a velocità non infinita doveva anche applicarsi agli elementi di questa realtà e alle reali influenze reciproche di questi elementi (e non

unicamente ai «fenomeni» intesi come «apparenze osservabili» «risultati di misurazioni»).

Su quest'ultimo punto – la localizzabilità degli elementi della realtà e la velocità finita delle influenze che essi esercitano gli uni sugli altri – si sa oggi che Einstein si è spinto troppo lontano. Come si è riportato sopra, la violazione delle disuguaglianze di Bell mostra che se si considera la nozione di realtà fisica dotata di senso e la si intende come riferentesi all'obiettivo della scienza fisica, questa realtà è allora non-separabile, il che vuol dire che vi si esercitano delle influenze a distanza che violano il principio di propagazione a velocità finita.

Come avrebbe reagito Einstein di fronte a questi risultati? Si può anche pensare che sarebbe passato alla tesi operazionista, ma sembra di gran lunga piú verosimile che avrebbe optato per una concezione realistica, con la quale avrebbe riconosciuto la realtà indipendente come non-separabile. Ogni volta che si evoca questa nozione, bisogna purtroppo ricordare che al momento attuale, malgrado i grandissimi sforzi profusi in questo senso da piú di mezzo secolo, non si dispone ancora di una teoria fisica pienamente soddisfacente e generale di una tale realtà.

Ma c'è un altro punto di vista. Come già è stato notato parlando di Kant, il fisico Heisenberg in particolare si è ispirato all'opera di questo filosofo per tentare di capire il problema dei rapporti dell'uomo con il mondo esterno, come lo pone la disciplina che lui stesso aveva in larga parte fondato. Ora, è facile, ma semplicistico, giustificare la scelta di una tale fonte di ispirazione con semplici argomenti psicologici (fondati ad esempio sull'appartenenza dei due autori a una stessa tradizione culturale). In realtà, la scelta in questione fu evidentemente motivata essenzialmente dal fatto che essa socchiude almeno qualche prospettiva di soluzione a un problema – quello della conciliazione tra meccanica quantistica e la nozione di oggetto reale – che sembra altrimenti insolubile o, in ogni caso, senza manifesta soluzione.

Ma Heisenberg ha modificato molto il kantismo per adattarlo alla situazione venutasi a creare. Si può dire, schematicamente, che ha fatto scendere l'a priori kantiano dal piedestallo. Cosí, ad esempio, l'innatismo delle nozioni di causalità, di tempo e di spazio – e quindi anche di località – non è piú in Heisenberg un dato trascendente posto una volta per tutte. Nella filosofia di questo autore tali concetti appaiono piuttosto come *acquisizioni*. Si tratta tuttavia di acquisizioni alle quali la loro stabilità e il loro ruolo essenziale nel padroneggiamento dei fenomeni da parte dell'uomo bastano a conferire attributi simili in pratica a quelli degli a priori kantiani. Infatti i concetti in questione, maturati senza dubbio nel corso del lungo combattimento con le cose in cui si è forgiato lo spirito umano, formano oggi l'impalcatura stessa del linguaggio. In particolare costituiscono la cornice inevitabile di ogni descrizione non ambigua degli strumenti di osservazione o di misura. Anche nella scienza è solo per loro tramite che, a detta di Bohr (che Heisenberg in questo segue da vicino), ci si può informare, gli uni con gli altri, di quel che si è fatto e di quello che si è appreso. Fondato su tali concetti, il linguaggio quotidiano – questo linguaggio di «vecchie parole» che, raffinato dall'aggiunta di parole tecniche, ha generato quello della fisica classica – è in verità il solo, almeno secondo questi due autori, che possa effettivamente svolgere un ruolo di comunicazione non ambigua. Se bisogna prestar loro fede, sarebbe dunque vano voler fondare i concetti della fisica classica – quelli che servono ad esempio per descrivere gli strumenti – su quelli della fisica quantistica, come funzioni d'onda, vettori di stato, ecc. Si è invece giustificati a fare il contrario: e siccome è proprio questo il modo di operare della fisica quantistica (che definisce le grandezze fisiche e i loro valori a partire dalle misure strumentali) risulta essa stessa giustificata nei suoi metodi e sollevata da ogni accusa d'incoerenza.

Non risulta per ciò meno vero che se si chiama «realtà» l'insieme dei fenomeni cosí descritti, una tale realtà non è indipendente dall'uomo: e che essa si avvicina in questo modo molto piú al reale kantiano che al mondo di Leibniz, alla sostanza di Spinoza o alla materia di Lenin. In modo piú preciso si vedono già delinearsi in Heisenberg alcuni elementi dell'idea, affinatasi poi in Prigogine, ad esempio, di una sorta di effetto circolare. I principî quantistici uniti a un principio di causalità che privilegia una direzione del tempo presiedono all'apparizione di quei fenomeni che sono i sistemi complessi; l'esistenza di questi ultimi presiede a sua volta alla comparsa degli esseri viventi, coronati dallo spirito umano; ma, a chiudere il cerchio, quest'ultimo deve proiettare le proprie categorie nelle cose, quando formula i principî quantistici o pone la causalità. Un'idea del genere è manifestamente incompatibile con l'ipotesi, secondo cui la scienza potrebbe legittimamente mirare a descrivere un ipotetico «reale in sé».

Ma se è difficile descrivere il reale in sé, è ancora piú difficile pretendere di eliminarne il concetto senza fare dei giochi di parole. Anche l'idea di effetto circolare qui sopra descritta diventerebbe allora, almeno pare, molto simile a quella di una qualche generazione spontanea dei fenomeni dal nulla, idea assai difficile da concepire. Si può cosí constatare che il concetto di una realtà indipendente resta presente in filigrana, anche nel pensiero dei fisici piú attenti all'epistemologia. Negli scritti di Heisenberg la nozione di *potenzialità* – presentata senza dubbio intenzionalmente in modo da lasciarla vaga – svolge indubbiamente un ruolo di questo tipo. Attualmente altri ricercatori, che si muovono a loro volta sulle tracce di Einstein, si sforzano, ancora una volta, d'integrare il concetto in questione alla fisica, dunque di renderlo netto e preciso. Non è sicuro che il loro tentativo risulti fecondo. Lo stato attuale di questa scienza è lontano dal rendere inverosimile la tesi di un realismo non fisico che, pur rispettando quanto c'è di legittimo nel concetto di una realtà *in sé*, fondamento delle regolarità osservabili, consideri una tale realtà in definitiva non descrivibile dalla fisica. Secondo questa tesi la scienza dovrebbe quindi limitare le proprie ambizioni alla pura descrizione dei fenomeni, non aspirare che all'obiettività debole e rinunziare a conoscere altro che strutture che hanno un'analogia parziale con la realtà in questione. In contropartita potrebbe legittimamente rivendicare – almeno a termine – l'esclusività

per quel che riguarda la descrizione appunto, dei fenomeni (cosa che, si noti bene, basta a mettere fuori gioco le magie).

Di fronte all'analogia cosí postulata tra le strutture della fisica e quelle della realtà indipendente, come di fronte alla spiegazione, che a quella è connessa, delle regolarità dei fenomeni, si constata che la tesi sopra enunciata prende evidentemente nei confronti del kantismo distanze tanto grandi quanto quelle che prende nei confronti dello spinozismo o del realismo ingenuo.

Non costituisce però altro che una tesi plausibile, la cui adozione o il cui rifiuto (a favore ad esempio della speranza di un possibile ritorno alla concezione realista di Einstein) dipendono in parte dal temperamento di ciascuno. A chi vuole aderire ad un realismo einsteiniano spetta tuttavia di provare senza cadere nella fantascienza che un tale realismo è possibile. Un tale compito, ancora una volta, si rivela difficile.

Se una conclusione è necessaria a questo articolo, si potrebbe proporre questa. Anche a prescindere dai contributi alle possibilità di azione dell'uomo – che richiedono una deontologia che andrebbe discussa a parte – l'apporto della fisica alla conoscenza resta essenziale. Questa scienza dimostra in particolare i seguenti fatti: i fenomeni dipendono da una descrizione razionale; questa descrizione è coerente; la sua vocazione a una perfetta universalità (dopo completamenti, estensioni e ritocchi, se necessario) sembra sulla buona strada per essere provata. Essa si fonda sull'oggettività *debole*. La sua interpretazione nei termini di un'attribuzione dei fenomeni a una realtà indipendente (oggettività forte) si rivela perlomeno difficile e ambigua. Ogni *realismo prossimo* – cioè ogni realismo che erige ad assoluto alcuni concetti della vita familiare e si fonda sull'attribuzione di *proprietà in sé*, calcate sui concetti in questione, a degli esseri molteplici e localizzati definibili senza riferimento all'uomo – tutto questo *realismo prossimo* sembra essere di conseguenza condannato irrimediabilmente dalla fisica, in quanto in definitiva costituisce un quadro inadeguato alla descrizione coerente dell'insieme dei fenomeni. [B. D'E.].

Einstein, A.
1953 *Einleitende Bemerkungen über Grundbegriffe*, in *Louis de Broglie physicien et penseur*, Michel, Paris.

Sia che venga caratterizzata rispetto ai propri oggetti di indagine sia al proprio **metodo**, la fisica pone e affronta le questioni di base della **conoscenza** della **natura**. Abbandonata l'idea di un'intuizione diretta delle essenze (cfr. **essere, esistenza, astratto/concreto**), essa sembra vertere sul mondo dei fenomeni (cfr. **fenomeno**): i procedimenti di osservazione e di misura si rivelano allora indispensabili. Ma la stessa natura dell'**ipotesi** in fisica e la struttura delle teorie (cfr. **teoria/modello**) mostrano che le **matematiche** intervengono a definire i concetti stessi (cfr. **concetto, idea**) che il fisico cerca di mettere in corrispondenza coi fenomeni osservati: le idee di fondo

della matematica (cfr. **strutture matematiche, calcolo, differenziale, funzioni, numero, probabilità, simmetria**) rendono possibile sia il passaggio da concezioni vaghe e qualitative a concettualizzazioni precise e quantitative (cfr. **qualità/quantità, approssimazione**), sia il controllo empirico delle teorie stesse (cfr. **empiria/esperienza, esperimento, previsione, verificabilità/falsificabilità**). Nello studio del **moto** come in quello della **materia** e dell'**energia**, nell'analisi dei processi reversibili e irreversibili (cfr. **entropia, equilibrio/squilibrio**), nelle spiegazioni riuscite dei fenomeni termici, elettromagnetici, ottici, ecc. (cfr. **conservazione/invarianza, gravitazione, luce**), nelle sue opposizioni ormai classiche (come **forza/campo**), nelle aree, infine, di ricerca avanzata (cfr. **atomo e molecola, particella, stato fisico, plasma**), la fisica non cessa di fornire dei paradigmi (cfr. **paradigma**) alla **ricerca** e di costituire un modello di **scienza**.

I grandi programmi di ricerca contemporanei, come la **relatività** e la fisica dei **quanti**, modificano le categorie di fondo (cfr. **categorie/categorizzazione**), quali **causa/effetto** (cfr. anche **determinato/indeterminato, caso/probabilità**), mutano le concezioni usuali dello spazio e del tempo (cfr. **spazio/tempo**), offrono una vasta gamma di soluzioni per classiche opposizioni (cfr. **continuo/discreto, locale/globale**), rimettono in discussione, infine, la natura della **spiegazione** scientifica e dello status delle leggi di natura (cfr. **legge**). Apparentemente divisa in specifiche discipline (cfr. **disciplina/discipline**), la fisica non sembra quindi rinunziare a una piú profonda unità, sia come punto di partenza per moderne **cosmologie**, sia come base teorica dello stesso processo tecnologico (cfr. **macchina**), sia infine nel rimandare ai problemi di fondo della conoscenza **del reale** (cfr. **soggetto/oggetto**) e alle grandi opzioni di natura filosofica (cfr. **coppie filosofiche, filosofia/filosofie**).

Forza/campo

1. *Le interazioni piú significative.*

Nella dicotomia 'forza/campo' è sicuramente il primo termine quello che, almeno nel lettore non specialista, suscita un'eco piú rilevata, come quello che, entrato largamente nell'uso corrente, e anzi da questo mutuato, possiede un significato in qualche misura riconducibile ad esperienze elementari. Non è dubbio infatti che nella mente di coloro che tra i primi si occuparono di scienza, dall'antichità a noi nota fino ai tempi della prima rivoluzione scientifica – e si pensi pure anche a Newton – il termine 'forza' racchiudesse un significato in prima istanza intuitivo e comunque legato a esperienze materiali umane – la forza muscolare – o naturali elementari – la forza del vento, ecc.

Non ci si propone qui in alcun modo di ripercorrere, neppure schematicamente, l'evolversi del significato di questo termine nel corso dei secoli, né di esaminare quali siano stati i suoi antecedenti semantici nella fisica di Aristotele e nella fisica medievale; esistono ottimi studi al riguardo, e qui si citeranno, come esempio, quelli di Hesse [1961] e Jammer [1957]. Si vuole soltanto evidenziare il ruolo svolto dal concetto *forza* in alcuni momenti rilevanti della storia della scienza degli ultimi tre secoli e dire come ad esso si sia affiancato, un po' alla volta e in modo talvolta alternativo talaltra complementare, fino a diventare oggi il polo dominante, il concetto *campo*.

Si comincerà con l'osservare che, nel concreto della ricerca e del pensiero dei fisici, le forze piú importanti non sono poi molte: se si escludono le forze legate ad esperienze elementari sopra menzionate che, pur fornendo un fondamentale paradigma psicologico, difficilmente costituiscono parte integrante di un esperimento scientifico, e quindi oggetto di accurata misurazione, la prima classe di forze che ha formato oggetto di studio sistematico è stata quella che comprende la forza di gravità; quella forza cioè che fa sí che i gravi cadano sulla superficie della Terra, o che il sistema solare e l'intero universo si muovano su orbite ben precise.

Un secondo tipo di forze di importanza cruciale è quello connesso con i fenomeni elettromagnetici; con quei fenomeni cioè, di cui si parlerà diffusamente anche piú oltre, che vanno dal mutuo attrarsi di due palline dotate di cariche di segno opposto e dall'orientarsi dell'ago della bussola, fino alla propagazione delle onde dette appunto elettromagnetiche, di cui la luce visibile, le onde radio, i raggi X non sono che esempi particolari.

V'è poi naturalmente una vasta classe di forze che agiscono a livello microscopico, atomico e nucleare, cui si accennerà nel seguito solo di sfuggita, dato anche che esse sono assai meno completamente inquadrate in una teoria fisica consistente di quanto non avvenga invece per i primi due tipi considerati.

Già in questa schematica classificazione è implicita l'osservazione che ad ogni tipo di forza corrisponde quello che viene usualmente chiamato il suo

range, ovvero il raggio d'azione entro il quale essa produce effetti «sensibili». Tipicamente, ad esempio, le forze gravitazionali sono quelle largamente dominanti sulle grandi distanze, interplanetarie, interstellari e intergalattiche, non nel senso che tra oggetti cosí lontani non vi siano altre interazioni – il fatto che noi *vediamo* la luce delle stelle significa che esiste interazione elettromagnetica tra noi e loro – ma nel senso che, dal punto di vista «quantitativo», le forze gravitazionali sono cosí dominanti che il moto di stelle e pianeti è con ottima approssimazione determinato soltanto da queste. In realtà anche le forze elettromagnetiche sono forze a lungo raggio, ma, poiché esistono cariche elettriche sia positive che negative (al contrario della massa, che può essere solo positiva), non si dà il caso nell'universo di grosse masse dotate di rilevante carica elettrica complessiva e quindi non vi sono rilevanti forze di natura coulombiana agenti sulle grandi distanze. Nel corto *range* d'altra parte, cioè a livello microscopico, le forze gravitazionali sono enormemente minoritarie rispetto alle forze elettromagnetiche e queste a loro volta sono largamente inferiori alle forze nucleari e subnucleari, cosí che, a seconda del tipo di fenomeni e quindi dell'ordine di distanze che si considerano, è lecito e conveniente tener conto solo di alcuni fra i tipi di forze che teoricamente agiscono.

In realtà questo parlare di «forze» non sembra presentare problemi finché il discorso viene tenuto a livello qualitativo e descrittivo, mentre le difficoltà si moltiplicano non appena si tenta di passare a un livello quantitativo ed esplicativo dei fenomeni, supponendo qui schematicamente di intendersi su ciò che la locuzione «spiegazione di un fenomeno» significhi.

2. *L'azione a distanza. Newton e la tradizione newtoniana.*

Come si vedrà, già in Newton esiste il problema: qual è la natura della gravità, cioè delle forze gravitazionali? Nei *Principia* egli non si pronunzia sulla questione anche se ne avverte l'esistenza e l'urgenza. Scrive infatti: «Userò le parole attrazione, impulso o propensione di qualcosa verso un centro, indifferentemente e promiscuamente l'una per l'altra; visto che queste forze sono considerate non fisicamente ma matematicamente. Per cui, il lettore si guardi dal credere che io con queste parole abbia voluto definire una specie o un modo d'azione o una causa o una ragione fisica, o che io, se per caso parlerò di centri che attirano o di centri muniti di forza, attribuisca le forze, in un senso reale e fisico, a centri (che sono soltanto punti matematici)» [1713, trad. it. pp. 100-1]; e piú avanti: «In generale assumo qui la parola *attrazione* per significare una qualsiasi tendenza dei corpi ad accostarsi l'uno all'altro;... Nello stesso senso generale assumo la parola *impulso*, in quanto in questo trattato esamino, come ho spiegato nelle definizioni, non le specie delle forze e le qualità fisiche, ma le quantità e le proporzioni matematiche» [*ibid.*, p. 339].

Rimaneva tuttavia il problema: in base a quale meccanismo si poteva concepire che una massa materiale situata in un punto dello spazio in un dato istante esercitasse un'*azione a distanza*, nello stesso istante, su qualsiasi altra

massa, vicina o lontana? Si trattava fondamentalmente della difficoltà di concepire una reciproca influenza che non avvenisse né per contatto diretto (quale si pensava agisse nell'urto), né per mediazione di alcun mezzo interposto (come, si direbbe oggi, nella propagazione del suono nell'aria). Difficoltà, come si comprende, tutt'altro che banale e che costituisce un po' il nocciolo della contraddizione di cui qui ci si occupa. Si precisa anzi subito che il punto di vista che ritiene del tutto concepibile e corretto parlare di azione a distanza, che vi sia o no un mezzo interposto, è quello che fa riferimento al primo termine di tale contraddizione concedendo ancora credibilità al concetto di forza come concetto base nello studio delle interazioni, mentre il ritenere che, in qualche senso, sia indispensabile un mezzo interposto che trasmetta l'influenza significa assumere come fondamentale il concetto di campo.

Non si accenna invece qui alla storia della critica al concetto di forza in quanto tale, storia che ci sembra meno interessante una volta acquisito il fatto che, data la difficoltà perfino di concepire una misura delle forze in sistemi meccanici non in equilibrio, il concetto di forza può essere a tutti gli effetti ritenuto un'abbreviazione di «massa per accelerazione», almeno nel senso (si veda anche oltre) che dire che un corpo P_1 esercita su un corpo P_2 una certa forza significa dire che è in grado di imprimergli una certa accelerazione (definibile e misurabile rigorosamente); come Ernst Mach [1883] ha mostrato, a ciò va aggiunto il fatto che si può convenientemente assegnare a ogni corpo una certa grandezza, la sua massa (inerziale), che ricopre consistentemente il ruolo di costante di proporzionalità nelle reciproche interazioni.

Tornando a Newton: l'insofferenza verso l'azione a distanza è già dei suoi contemporanei (per non risalire piú indietro): «In confronto all'idea meccanicistica dell'urto come l'unica causa di mutamento dello stato di movimento, v'è un sapore spirituale o animistico nell'idea di una forza incomprensibile che agisce a distanza e che non sembra in disaccordo con una filosofia della natura antimaterialistica» [Dijksterhuis 1950, trad. it. p. 652]. Oggi le riserve di cartesiani e leibniziani verso le «qualità incorporee e inesplicabili» trovano ancora un'inaspettata eco: «Descartes, con i suoi vortici e i suoi atomi uncinati, spiegava tutto e non calcolava nulla; Newton con la legge di gravitazione in $1/r^2$ calcolava tutto e non spiegava nulla. La storia ha dato ragione a Newton e ha relegato le costruzioni cartesiane al rango di fantasie gratuite e di souvenirs da museo. Certo, il punto di vista newtoniano si giustifica pienamente sotto il profilo dell'efficacia, della possibilità di previsione, quindi d'azione, sui fenomeni... Ma non sono affatto sicuro che in un universo in cui tutti i fenomeni sono retti da uno schema matematicamente coerente, sprovvisto però di contenuto intuitivo, lo spirito umano sia pienamente soddisfatto. Non saremmo in tal caso in piena magia?... Infatti, nel 1968, non abbiamo meno ragioni di stupirci della caduta della mela di quante non ne avesse Newton» [Thom 1972, p. 21].

Peraltro lo stesso Newton in piú luoghi oppone resistenza a ogni tentativo di attribuirgli opinioni del genere! Valga per tutti il seguente passo di una lettera a Bentley: «È inconcepibile che materia bruta e inanimata possa, senza la mediazione di qualcos'altro che non è materiale, operare e influire su altra materia senza reciproco contatto... Che la gravità possa essere innata, intrinseca ed essenziale alla materia, cosicché un corpo possa agire su un altro a distanza attraverso il vuoto, senza la mediazione di qualcos'altro, in virtú del quale e mediante il quale la loro azione e forza possa essere trasmessa dall'uno all'altro, è per me una tale assurdità che io credo che nessun uomo capace di pensare in modo coerente in materia di filosofia possa accettarla» [citato in Hall 1954, trad. it. p. 254]. Qui come altrove, tipicamente nelle *Queries*, poste come appendice all'*Opticks* [1704], dove Newton propone alcuni meccanismi che rendano conto dell'attrazione gravitazionale, appare chiaro quanto forte fosse in lui la ripugnanza per una teoria basata sull'azione a distanza. Va d'altra parte rilevato che, visto anche l'atteggiamento agnostico che mantenne a questo riguardo nei *Principia*, ciò che nella conoscenza comune è sopravvissuto dell'opera sua, almeno per quanto riguarda la teoria della gravitazione, è senz'altro la bruta formula

$$\mathbf{F} = G\frac{m_1 m_2}{r^2}\mathbf{u}$$

per rappresentare la forza **F** che la massa m_1 esercita sulla massa m_2 posta a distanza r (con G costante che dipende dalle unità di misura usate e **u** versore orientato da m_2 a m_1), e non certo le complicate teorie ipotizzate nelle *Queries* e altrove per «spiegare» la gravità; tanto è vero ciò, che in un pionieristico articolo di Wheeler e Feynman, pur per altri versi assai informato e importante come si dirà meglio piú oltre, è dato reperire la seguente affermazione: «Era il 19 marzo 1845 quando Gauss descrisse il concetto di un'azione a distanza che si propaga con velocità finita, naturale generalizzazione all'elettrodinamica del punto di vista sulla forza cosí fruttuosamente utilizzato da Newton e dai suoi successori»! [1945, p. 157].

3. *I fluidi sottili.*

Nel Settecento la questione fu agitata da filosofi e scienziati (se qualche senso questa distinzione conserva per quel secolo) coinvolgendo Leibniz, Boscovich, gli empiristi inglesi e altri, depositari taluni dell'eredità cartesiana, piú newtoniani dello stesso Newton altri. Elemento comune di varie teorie emergenti era comunque l'esistenza di un fluido imponderabile, altamente sottile e penetrante, tanto da pervadere con maggiore o minore densità tutto l'universo, il cosiddetto *etere*; cosí come in altri campi della fisica altri fluidi analoghi venivano ipotizzati con lo scopo di «spiegare» vari fenomeni: si pensi al flogisto per le reazioni chimiche, al fluido elettrico per i fenomeni elettrostatici e al calorico per gli scambi di calore della termodinamica. In qualche modo, a parte le distinzioni pur rilevanti fra le teorie che ipotizzavano questi fluidi, tutti rispondevano all'esigenza di fornire un modello esplicativo intelligibile a fenomeni di cui altrimenti, almeno allora, si sarebbe dato poco piú

d'una descrizione qualitativa. Tali modelli erano dunque piú che mere analogie: svolgevano in realtà una doppia funzione, di concettualizzazione e di calcolo.

Quale fu l'evoluzione di questi tipi di modelli nell'Ottocento? Alcuni di essi cessarono semplicemente di esistere (il flogisto, il calorico, il fluido elettrico) sostituiti in qualche misura da spiegazioni microscopiche piú convincenti; altri, come l'etere, persero un po' alla volta consistenza. Si vuole dire con questo che anche se per tutto l'Ottocento non si cessò mai di credere nell'etere, tuttavia si pensò sempre meno a come era fatto e alla ragione per cui, essendo fatto cosí, poteva fare da supporto alla propagazione delle onde elettromagnetiche. L'etere svolgeva dunque, sempre di piú, il ruolo di giustificare la possibilità di tale propagazione, che avveniva – ciò era l'importante – secondo leggi ben precise, le equazioni di Maxwell. Anche a questa affermazione va naturalmente attribuito un significato di valutazione media di linea di tendenza: non mancarono certamente irriducibili meccanicisti sostenitori della necessità di modelli completamente esplicativi. Valgano d'esempio le seguenti parole di Lord Kelvin: «Innanzitutto non dobbiamo ascoltare alcuno di quei suggerimenti che tendono a farci considerare l'etere luminifero come una semplice costruzione mentale. Io credo che vi sia della materia reale tra noi e le stelle piú remote, e che la luce consista di movimenti reali di questa materia, movimenti proprio come quelli descritti da Fresnel e da Young, movimenti consistenti in vibrazioni trasversali» [1884, pp. 8-9].

L'azione a distanza, che pure aveva trovato i suoi sostenitori in varie epoche – spesso con diverse motivazioni – da Guglielmo di Ockham (contro la stragrande maggioranza degli scolastici) a Keplero, da Boscovich agli empiristi inglesi del Settecento, perde un po' alla volta terreno a favore della teoria di campo che, mentre in un primo tempo serve appunto a far superare le difficoltà filosofiche e «logiche» connesse con quella, in un secondo tempo conosce un'evoluzione formale cosí articolata che ne fa uno degli strumenti teorici piú massicciamente sviluppato fino ai nostri giorni.

4. *Il campo come paradigma dominante: Maxwell.*

Si riprende ora, per ampliare un poco i termini della discussione, il filo di quest'evoluzione al tempo e nell'opera di Maxwell. I suoi contributi alla costruzione di una teoria dell'elettromagnetismo sono, com'è noto, del tutto cruciali e ancor oggi le equazioni basilari della teoria sono appunto note come equazioni di Maxwell. Quattro sono gli scritti in cui egli elaborò complessivamente la teoria, ma dai primi due agli ultimi due vi è una notevole differenza di accento per quel che riguarda il famoso problema dell'esistenza e della costituzione dell'etere. Nel secondo di tali scritti, in particolare, cioè in *On physical lines of force* [1861], Maxwell dà veste quantitativa al suo modello di etere: si tratta di un fluido sottile ed elastico costituito da un sistema di piccoli vortici («vortici molecolari») i cui assi di rotazione sono orientati come le linee di forza magnetica. Si considerino ad esempio le linee di forza (ampiamente descritte e studiate da Michael Faraday, sulla cui opera sperimentale i lavori di Maxwell erano largamente fondati) che congiungono i due poli, nord e sud, di un magnete, o di due magneti distinti; fra essi si esercita una certa attrazione, mentre un filo percorso da corrente elettrica disposto in direzione ortogonale a tali linee è soggetto a una forza in direzione ortogonale a entrambe le precedenti e la cui intensità dipende da quelle della corrente e del magnete, oltre che dalla configurazione geometrica del sistema; si veda ora come tutto ciò può venire spiegato mediante il modellino dei vortici elementari di Maxwell, il cui asse si suppone dunque disposto come le linee di forza che congiungono i due poli; tali vortici presentano, in virtú del moto rotatorio che li anima, e della conseguente forza centrifuga, una tendenza a «contrarsi», cioè a diminuire la propria lunghezza ingrossandosi al centro, dando cosí luogo all'azione di attrazione tra le estremità dei vortici e a una parallela azione sulle correnti elettriche. Consideriamo, per semplificare, la seguente schematizzazione (cfr. fig. 1*a*): supponiamo che un polo nord magnetico P sia posto in una regione nella quale le linee di forza sono parallele e uniformi, dirette verso l'alto; le linee di forza provenienti da P (tratteggiate in figura) saranno disposte in maniera radiale con centro P; la linea diretta verso l'alto si sovrappone *sommandosi* a quella preesistente; ciò corrisponde a vortici la cui velocità angolare aumenta, cosí che le estremità tendono a contrarsi; il contrario avviene in direzione opposta, cosí che, complessivamente, P viene attirato verso l'alto. Analogamente, nella stessa regione, supponiamo di considerare un filo percorso da corrente in direzione ortogonale al piano del disegno e con verso uscente verso chi legge (cfr. fig. 1*b*). Sia F l'intersezione del filo col piano del disegno; le linee di forza generate dalla corrente sono disposte circolarmente come in figura, cosí che a destra di F il loro effetto si somma a quello delle linee preesistenti e a sinistra si sottrae: i vortici a destra aumentano in velocità e aumenta quindi la loro tendenza ad allargarsi, mentre quelli a sinistra di F perdono velocità e quindi tensione trasversale;

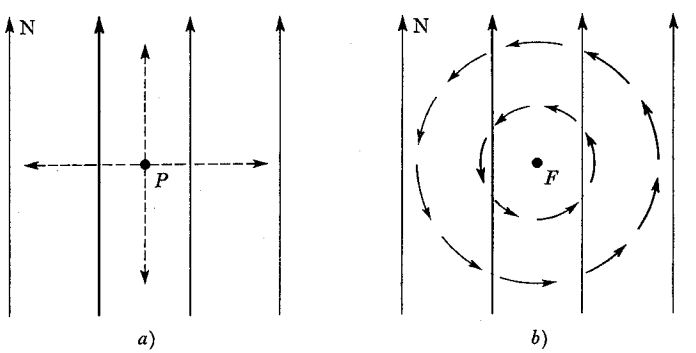


Figura 1.

Polo nord magnetico P (*a*) e filo percorso da corrente (*b*) in un campo di forze magnetiche uniforme.

i vortici di destra hanno quindi il sopravvento e spingono F verso sinistra, come appunto sperimentalmente accade. Per superare la difficoltà, consistente nel fatto che per due vortici adiacenti e rotanti nel medesimo senso le falde che si toccano si muovono in versi opposti, Maxwell ipotizzò l'esistenza tra vortice e vortice di lunghe file di minuscole particelle, con l'ufficio, esattamente, di cuscinetti a sfere, adatte a rotolare senza strisciare sulle falde dei vortici; tali particelle vennero supposte da Maxwell essere i costituenti elementari dell'elettricità. Un tale modello si prestava, come si è ricordato, a un'analisi quantitativa sufficientemente accurata da produrre le principali formule dell'elettromagnetismo; di fatto Maxwell arrivò a mostrare come su questa base si poteva giustificare e comprendere la propagazione, attraverso questo mezzo etereo fatto di vortici e palline, della luce in quanto fenomeno ondulatorio trasversale viaggiante con velocità dipendente, com'è del tutto confermato dall'esperienza, dalle proprietà dielettriche e magnetiche del mezzo. Le famose equazioni di Maxwell potevano dunque venir dedotte da un modello meccanico di etere fluido ed elastico. Se il modello presenta non poche forzature e difficoltà al nostro occhio, costituiva tuttavia a quel tempo uno sforzo notevole di spiegazione dei fenomeni elettromagnetici.

Tra le difficoltà non si deve dimenticare la seguente, che si applica sia a questo modello che ad altri analoghi nei quali le interazioni sono supposte essere mediate e trasmesse da qualche mezzo meccanico; se questo mezzo possiede una struttura discreta, per esempio è costituito da tante palline, come si trasmette la perturbazione da ogni pallina a quella immediatamente attigua, che, per varie ragioni, è difficile immaginare continuamente a contatto con la precedente? In altre parole l'azione a distanza, scacciata a livello macroscopico, si ripropone a livello microscopico. Difficoltà difficilmente superabile, a quei tempi, nell'ambito di una teoria atomistica (fu questo del resto uno degli scogli che si frappose al successo della teoria atomistica della materia).

Si legga ora invece questo brano, che val la pena di citare letteralmente, tratto dai commenti finali della parte III del terzo scritto di Maxwell sull'elettromagnetismo, *A dynamical theory of the electromagnetic field* [1864], che contiene tutta la materia essenziale che verrà poi ampliata e organizzata nel *Treatise* [1873]: «In un'occasione precedente [in *On physical lines of force*] ho cercato di descrivere un particolare tipo di moto e un particolare sistema di tensioni, disposti in modo da render conto dei fenomeni. In questo scritto evito ogni ipotesi di questo tipo; ed usando parole come momento elettrico ed elasticità elettrica in riferimento ai noti fenomeni dell'induzione di corrente e della polarizzazione del dielettrico, desidero soltanto indicare alla mente del lettore fenomeni meccanici che gli saranno d'aiuto nel comprendere i fenomeni elettrici. Tutte queste espressioni in questo scritto vanno considerate come illustrative, non come esplicative. Quando parlo, tuttavia, dell'energia del campo, desidero venir inteso letteralmente. Tutta l'energia è dello stesso tipo dell'energia meccanica (*the same as mechanical energy*), sia che esista sotto forma di moto, o di elasticità, o in qualsiasi altra forma. L'energia nei fenomeni elettromagnetici è energia meccanica. Il solo problema è: dove essa ha

sede?... Nella nostra teoria risiede nel campo elettromagnetico, nello spazio circostante i corpi elettrizzati e magnetizzati, cosí come entro questi stessi corpi» [1864, pp. 487-88]. E da qui in poi la descrizione dei fenomeni elettromagnetici viene condotta seccamente in termini di campo.

In queste poche frasi, e piú ancora nella struttura di tutto il lavoro, cosí come nel *Treatise*, è contenuto un notevole cambiamento di prospettiva rispetto ai primi due scritti: è qui individuabile uno degli episodi rilevanti di una rivoluzione epistemologica di grande portata, che presenta diversi e problematici aspetti, ma che si può comunque cogliere nel passaggio da un ideale esplicativo, meccanico, modellistico in senso realistico, che ispirò la maggior parte della ricerca fisica dell'Ottocento, a un ideale piú astratto e formale, modellistico in senso logico, e comunque piú matematizzante, che ha invece prevalso grosso modo dai primi anni del secolo fino ai giorni nostri.

Il brano di Maxwell citato illustra in particolare un modo nuovo di porre il problema della trasmissione a distanza di perturbazioni elettromagnetiche e avvia una soluzione che da quel momento in poi fu sempre piú dominante. Maxwell afferma ancora che la propagazione delle onde elettromagnetiche avviene in un mezzo dotato di un funzionamento complicato e comunque a noi ignoto, del quale è tuttavia sufficiente quella descrizione potente e complessiva, che non entra nel dettaglio sconosciuto del meccanicismo di trasmissione, fornita dalle quattro equazioni differenziali che portano il suo nome. L'accenno, contenuto nelle parole di Maxwell, alla questione dell'energia è assai significativo ed è legato ad una delle principali difficoltà che si frappongono a una teoria di azione a distanza; difficoltà che, espressa nel modo piú semplice possibile, è la seguente: se una perturbazione elettromagnetica parte da un punto A e arriva in un punto B dopo un certo tempo (è questo un elemento essenziale: l'interazione si propaga con velocità finita, non istantaneamente), noi possiamo constatare l'esistenza in A di una certa energia ad un tempo t_A e in B della stessa energia ad un istante successivo t_B; negli istanti compresi tra t_A e t_B dov'è l'energia? Le innumeri conferme che il principio di conservazione dell'energia ha avuto nella fisica rendono difficile asserire che in quell'intervallo di tempo l'energia è sparita; ecco quindi che supporre invece che lo spazio intero sia sede di un *campo*, di cui non si può dire con troppa precisione *cos'è* (per il senso che ha qui questa domanda), ma di cui si possono dire chiaramente tutte le proprietà osservabili, risolve la difficoltà, visto che è possibile in maniera consistente assegnare a ogni valore del campo, mediante opportune prescrizioni formali, valori precisi dell'energia (e del momento) ad esso associati.

5. *Aspetti formali della questione.*

È necessario a questo punto accennare a qualche aspetto piú tecnico dei discorsi fin qui fatti, senza il quale non sarebbe possibile apprezzare sufficientemente in che consiste la differenza tra teorie di campo e azione a distanza.

L'equazione di Newton, fondamento della dinamica classica, è un'equazione differenziale ordinaria del second'ordine, nella quale l'unica variabile rispetto alla quale si deriva è il tempo; se $\mathbf{x}(t)$ rappresenta la posizione della particella di massa m al tempo t e $\mathbf{F}(\mathbf{x},t)$ la forza che agisce sulla particella nel punto $\mathbf{x}$ all'istante t, l'equazione di Newton si scrive infatti

$$(1) \qquad \mathbf{F}(\mathbf{x},t)=m\frac{d^2\mathbf{x}}{dt^2}(t),$$

dove $\mathbf{F}$ è supposto assegnato, cosí come m, e l'incognita è la funzione $\mathbf{x}$; non deve trarre in inganno il fatto che, nella forma di Hamilton, le equazioni della dinamica si scrivano

$$\frac{dq}{dt}=\frac{\partial H}{\partial p} \qquad \frac{dp}{dt}=-\frac{\partial H}{\partial q},$$

cioè con derivate *parziali* rispetto a p e a q della funzione H: questa infatti è supposta nota mentre le funzioni incognite sono p e q, funzioni del tempo solamente, che definiscono la traiettoria nello spazio delle fasi. Per le equazioni differenziali di questo tipo, dette ordinarie, è noto che, sotto opportune condizioni di regolarità, esiste una classe di soluzioni, al cui interno si individua la soluzione desiderata pur di fissare una o piú (a seconda del grado dell'equazione) costanti arbitrarie, il cui valore è usualmente dettato, in fisica, dai dati particolari del problema che si sta considerando. Tipicamente, nel caso di una particella soggetta a forze, la conoscenza della posizione e della, velocità ad un istante, che può essere ad esempio quello iniziale, individua fra tutte le soluzioni dell'equazione (1), quella adeguata al problema. Se questo procedimento viene seguito nel caso della teoria della gravitazione o dell'elettromagnetismo, occorre sapere, nel caso ovviamente di presenza di due o piú particelle, qual è la forza che una di esse nella posizione $\mathbf{x}$ esercita su un'altra che si trovi nella posizione $\mathbf{y}$, in virtú semplicemente della loro presenza contemporanea, in altre parole si ragiona in termini di azione a distanza tra corpi, indipendentemente da *come* quest'azione eventualmente si propaghi e, nel caso, con che velocità. L'evoluzione da questa concezione a una concezione campistica che, come si è accennato, avviene sulla base di esigenze esplicative profonde, anche se talvolta ingenue, dei fenomeni, viene piú spesso introdotta soprattutto nei manuali piú elementari e meno attenti, nel modo seguente: si consideri, per fissare le idee, il caso della gravitazione; una particella di massa m_1 situata in $\mathbf{x}$ fa sí che, se una seconda particella di massa m_2 si trova in $\mathbf{y}$, quest'ultima sia soggetta a una forza pari a

$$\mathbf{F}_{12}=G\frac{m_1m_2}{|\mathbf{x}-\mathbf{y}|^3}(\mathbf{x}-\mathbf{y}),$$

e ciò vale per ogni punto $\mathbf{y}$ distinto da $\mathbf{x}$; allora si può *descrivere* questa situazione dicendo che la presenza di m_1 in $\mathbf{x}$ crea in tutto lo spazio circostante un campo di forze $\mathbf{F}$; il che significa a questo livello soltanto che una particella di massa m posta in $\mathbf{y}$ risente di una forza $\mathbf{F}(\mathbf{y})$, ovvero, in termini un po' piú avveduti rispetto alle critiche al concetto di forza cui si accennava, risente di un'accelerazione pari a $(1/m)\mathbf{F}(\mathbf{y})$. E fin qui il concetto di campo non sarebbe altro che un comodo ausilio matematico per descrivere in realtà l'azione a distanza. *Esiste dunque*, nello spazio circostante m_1, un campo di forze $\mathbf{F}$. Quest'ultima affermazione del nostro supposto manuale porta bruscamente il campo di forze dal livello logico al livello ontologico. La spinta a questo passaggio viene beninteso accentuata da alcune aporie connesse con la teoria dell'azione a distanza, aporie di cui si è già detto e si dirà ancora, ma ciò non toglie che si tratti di un passaggio a questo livello del tutto ingiustificato. Esso può certamente venir compiuto, sulla base di motivazioni del tutto diverse, pur di esserne coscienti e di accettarne le implicazioni, senza nascondersi che cosí facendo si eliminano sí alcune difficoltà della teoria basata sull'azione a distanza, ma non si rimane certamente privi di difficoltà concettuali sulla natura del campo stesso.

Tra le aporie che per prime si presentarono a una formulazione ingenua dell'elettromagnetismo mediante una teoria di azione a distanza, ve n'è una che è opportuno accennare subito, in quanto fu il primo segnale di contraddizione dell'elettromagnetismo maxwelliano con la meccanica newtoniana. Si considerino due particelle dotate di cariche elettriche q_1 e q_2 e di velocità $\mathbf{v}_1$ e $\mathbf{v}_2$ in un fissato sistema di riferimento; se si scrive la forza $\mathbf{F}_{12}$ che q_1 esercita su q_2 e $\mathbf{F}_{21}$ che q_2 esercita su q_1, si ottiene dopo alcune semplificazioni

$$\mathbf{F}_{ij}=\frac{q_iq_j}{c^2r^3}[\mathbf{v}_i(\mathbf{v}_j\cdot(\mathbf{x}_j-\mathbf{x}_i))-(\mathbf{x}_j-\mathbf{x}_i)(\mathbf{v}_i\cdot\mathbf{v}_j)] \qquad i,j=\begin{cases}1,\ 2\\ 2,\ 1\end{cases}$$

dove c indica la velocità della luce nel vuoto e $r=|\mathbf{x}_2-\mathbf{x}_1|$; questa espressione è manifestamente non antisimmetrica negli indici i e j, cioè non è vero che $\mathbf{F}_{12}=-\mathbf{F}_{21}$ cosí che $\mathbf{F}_{12}+\mathbf{F}_{21}\neq 0$ in generale e anzi tale somma è nulla in alcuni sistemi di riferimento (per esempio quelli in cui $\mathbf{v}_1=0$ oppure $\mathbf{v}_2=0$) e non in altri; viene cosí contraddetto il terzo principio della dinamica, o se si vuole la legge di conservazione della quantità di moto di un sistema isolato.

In parole estremamente semplici, la teoria maxwelliana del campo elettromagnetico ricupera la validità di tale principio di conservazione conferendo all'entità «campo elettromagnetico» una realtà fisica concreta e attribuendogli di conseguenza una certa energia e una certa quantità di moto; inoltre il tutto deve venir riconsiderato alla luce della teoria della relatività speciale: infatti la teoria maxwelliana che, come si è visto, non è galileianamente invariante, era *ante litteram*, al suo sorgere, intrinsecamente relativistica.

Si vuole ora esaminare piú da vicino cos'è un'equazione di campo e in che cosa differisce da un'equazione che descrive forze meccaniche esercitate a distanza. Nella teoria maxwelliana, che qui non è certo possibile esporre neppure fuggevolmente, il campo elettromagnetico è descritto da quattro campi vettoriali (meglio sarebbe dire da due campi tensoriali antisimmetrici, ma ci si attiene qui al linguaggio vettoriale, forse piú familiare); uno di questi, indicato usualmente con la lettera $\mathbf{D}$, rappresenta il vettore spostamento elet-

trico ed ha il significato di rappresentare nello spazio vuoto il campo elettrico puro e semplice e, all'interno di un dielettrico, di tener conto della polarizzazione di quest'ultimo; se ρ è la densità di carica elettrica in ogni punto dello spazio (vuoto o no), **D** obbedisce, punto per punto, all'equazione (una delle quattro equazioni di Maxwell) $\operatorname{div}\mathbf{D}=4\pi\rho$ che lega dunque tale campo alla distribuzione di carica; si tratta in questo caso di un'equazione differenziale non piú ordinaria, ma alle derivate parziali: **D** è una funzione, a valori vettoriali, dei punti dello spazio, oltre che del tempo, e oltre alla derivata rispetto a quest'ultimo (che in questa specifica equazione non compare esplicitamente) appaiono le derivate di **D** rispetto alle coordinate spaziali (si ricordi che $\operatorname{div}\mathbf{D}:=\frac{\partial D_x}{\partial x}+\frac{\partial D_y}{\partial y}+\frac{\partial D_z}{\partial z}$); ciò significa che una tale equazione correla il comportamento del campo in un punto al comportamento in un punto vicino: è cioè un'equazione di tipo *locale*. Analogamente, se **H** è il campo vettoriale che rappresenta l'intensità di campo magnetico, la quarta equazione di Maxwell si scrive:

$$\operatorname{rot}\mathbf{H}-\frac{1}{c}\frac{\partial\mathbf{D}}{\partial t}=4\pi\mathbf{J}$$

dove **J** è il vettore densità di corrente; qui compaiono contemporaneamente derivate parziali rispetto a t e alle coordinate, e ciò ancora una volta significa che l'equazione correla i valori del campo in un punto e ad un istante con i valori in punti e istanti vicini; il campo, con la sua energia e il suo momento, media l'interazione tra corpi carichi, ha un'esistenza autonoma, trasporta una perturbazione da un punto ai punti vicini in un tempo finito. Scritte le equazioni per il campo elettromagnetico (rimangono altre due equazioni di Maxwell, che determinano proprietà del solo campo, senza relazioni con le distribuzioni di cariche e correnti, e sono $\operatorname{div}\mathbf{B}=0$ e $\operatorname{rot}\mathbf{E}-(1/c)(\partial\mathbf{B}/\partial t)=0$, essendo **E** e **B** rispettivamente il campo elettrico e l'induzione magnetica), occorre scrivere anche le equazioni del moto per i corpi materiali provvisti di carica che in tale campo sono immersi. Una difficoltà tipica a questo punto è che, poiché il campo complessivo è generato dai corpi carichi stessi, nei punti dello spazio da questi occupati esso non è definito (si vede ad esempio dalle leggi di Coulomb e di Biot-Savart che esso in tali punti ha limite infinito); occorre allora, per scrivere le equazioni di moto di una particella carica nel campo, prescindere dal campo da essa stessa generato e tener conto soltanto di quello generato dal resto delle cariche presenti. Questa distinzione, semplice in casi elementari, diventa problematica nel caso generale di cariche in moto con velocità non trascurabile rispetto a c, e non è univocamente determinata dalle equazioni di campo. Si ricorda infatti che le due equazioni di Maxwell indipendenti dalle sorgenti permettono di far derivare i campi, matematicamente parlando, da due altri campi, uno scalare φ e uno vettoriale **A**, dipendenti dalla distribuzione di cariche e correnti, mediante le formule

$$\mathbf{B}=\operatorname{rot}\mathbf{A}\qquad \mathbf{E}=-\operatorname{grad}\varphi-\frac{1}{c}\frac{\partial\mathbf{A}}{\partial t};$$

le altre due equazioni di Maxwell determinano la dipendenza di φ e **A** da ρ e **J**; ed è quest'ultima dipendenza che non è univocamente determinata, coerentemente col fatto che i potenziali elettromagnetici φ e **A**, che per sé non hanno significato fisico alcuno essendo soltanto un formidabile ausilio matematico, non sono univocamente determinati dalle equazioni di campo: ciò è una peculiare caratteristica del fatto che si tratta di equazioni alle derivate parziali, le cui soluzioni dipendono in generale dalla scelta di una funzione arbitraria in una certa classe di regolarità. Non è, in particolare, univocamente determinato dalle equazioni di Maxwell il potenziale elettromagnetico generato da un'assegnata distribuzione di cariche $\rho(\mathbf{x},t)$ e correnti $\mathbf{J}(\mathbf{x},t)$: per essi sono possibili almeno le due seguenti soluzioni:

$$(2a)\qquad \begin{aligned}\varphi^{(\mathrm{rit})}(\mathbf{x},t)&=\int_{\mathbf{R}^3}\frac{1}{|\mathbf{x}-\mathbf{y}|}\,\rho\left(\mathbf{y},t-\frac{|\mathbf{x}-\mathbf{y}|}{c}\right)d\mathbf{y}\\ \mathbf{A}^{(\mathrm{rit})}(\mathbf{x},t)&=\int_{\mathbf{R}^3}\frac{1}{|\mathbf{x}-\mathbf{y}|}\,\mathbf{J}\left(\mathbf{y},t-\frac{|\mathbf{x}-\mathbf{y}|}{c}\right)d\mathbf{y};\end{aligned}$$

questi vengono detti potenziali ritardati poiché, come si vede dall'integrando di entrambe le formule, il potenziale viene calcolato tenendo conto del contributo delle sorgenti (ρ e **J**) che sono collocate nel generico punto **y**, ma non all'istante t, ma ad un istante $t-(|\mathbf{x}-\mathbf{y}|/c)$ che è quello nel quale deve partire un segnale elettromagnetico da **y** per raggiungere **x** all'istante t; ma esiste anche la soluzione

$$(2b)\qquad \begin{aligned}\varphi^{(\mathrm{av})}(\mathbf{x},t)&=\int_{\mathbf{R}^3}\frac{1}{|\mathbf{x}-\mathbf{y}|}\,\rho\left(\mathbf{y},t+\frac{|\mathbf{x}-\mathbf{y}|}{c}\right)d\mathbf{y}\\ \mathbf{A}^{(\mathrm{av})}(\mathbf{x},t)&=\int_{\mathbf{R}^3}\frac{1}{|\mathbf{x}-\mathbf{y}|}\,\mathbf{J}\left(\mathbf{y},t+\frac{|\mathbf{x}-\mathbf{y}|}{c}\right)d\mathbf{y}\end{aligned}$$

formata dai cosiddetti potenziali avanzati (o anticipati) in quanto tengono conto della distribuzione esistente in **y** non ad un istante precedente t, ma ad uno successivo a t e dello stesso intervallo. La contraddizione di quest'ultima prescrizione di calcolo per i potenziali con un intuitivo principio di causalità (l'effetto deve precedere la causa) è in realtà piú apparente che reale e può venir superata mediante una formulazione piú accurata e soddisfacente del principio di causalità. Una soluzione anzi che presenta alcuni vantaggi è quella fornita dalla media aritmetica delle due precedenti:

$$(2c)\qquad \begin{aligned}\varphi&=\frac{1}{2}(\varphi^{(\mathrm{rit})}+\varphi^{(\mathrm{av})})\\ \mathbf{A}&=\frac{1}{2}(\mathbf{A}^{(\mathrm{rit})}+\mathbf{A}^{(\mathrm{av})});\end{aligned}$$

è di essa che si servono alcuni tentativi moderni di fare dell'elettromagnetismo una teoria dell'azione a distanza, segnatamente la formulazione di Schwarzschild e Fokker cui si farà cenno oltre.

6. *Il principio di Mach.*

Se a questo punto si vogliono esaminare alcuni punti di vista piú moderni su tutta la questione e dare un cenno dei tentativi piú recenti di formulare una teoria delle interazioni gravitazionali ed elettromagnetiche mediante il concetto di azione a distanza, si deve partire da alcune riflessioni cruciali con le quali Ernst Mach portò, dopo quasi due secoli di newtonianesimo piú o meno indiscusso, nuove e stimolanti idee nell'edificio della meccanica classica. Riconosciuto che strutture portanti di tale edificio sono alcuni sistemi di riferimento privilegiati, quelli detti inerziali, nei quali valgono i primi due principî della meccanica (detti di Galileo e di Newton rispettivamente), egli criticò con grande perspicuità l'esperimento del secchio ruotante con il quale Newton pensava di distinguere in maniera *assoluta* i riferimenti inerziali da quelli non inerziali. Il punto di vista di Newton era semplicemente il seguente: se colloco un secchio contenente dell'acqua in quiete in un sistema di riferimento inerziale, la superficie del liquido rimane piana; se lo pongo in un sistema di riferimento ruotante (rispetto a un sistema inerziale) la superficie del liquido assume una definita concavità, dipendente dalla velocità di rotazione; ciò permette dunque di distinguere in maniera assoluta, cioè dall'interno dello stesso sistema, se esso è inerziale o no. La critica di Mach, fondata non su argomentazioni quantitative ma su un'ipotesi qualitativa globale, fu che sia il fatto che i sistemi inerziali fossero di fatto quelli in quiete o moto rettilineo uniforme rispetto alle stelle fisse, sia l'incurvarsi o meno della superficie dell'acqua nel secchio, sono frutto della distribuzione di masse nell'universo; in altre parole, le masse delle stelle e di tutti i corpi materiali esistenti, che sono, per cosí dire, lontani ma tanti, sono ciò che realmente determina cos'è, in un qualsiasi punto dello spazio, un sistema inerziale. Va da sé dunque, nell'ipotesi di Mach, che se il secchio fosse il solo corpo materiale esistente nell'universo, poco senso avrebbe distinguere un suo stato di rotazione da uno di quiete e l'esperienza dell'osservazione della superficie dell'acqua in esso contenuta potrebbe mutare radicalmente e perdere senso.

Quest'ipotesi, nota da allora in poi come «principio di Mach», ebbe grande influenza sullo stesso Einstein ed ebbe probabilmente un notevole peso nella costruzione einsteiniana della teoria della relatività generale. In realtà, malgrado le speranze dello stesso Einstein e di altri, non si riuscí mai a ricavare il principio di Mach dalla relatività generale, e anzi si scoprí che in un particolare modello di universo [cfr. Gödel 1949], all'interno del formalismo della relatività generale, il principio di Mach non valeva, nel senso che esistono, in tale modello, sistemi di riferimento localmente inerziali, rispetto ai quali le masse lontane ruotano (*vorticità* dell'universo). Ciò significa quanto meno, dal nostro punto di vista, che la natura specifica dell'interazione gravitazionale tra masse lontane, inquadrata dalla relatività in uno schema complessivo che fa dipendere la geometria dell'intero universo (e quindi il modo di calcolare angoli e distanze) dalla distribuzione dei corpi in esso, è ancora lontana

dall'essere definitivamente chiarita. Una proposta di soluzione relativamente recente [Hoyle e Narlikar 1963] è quella di ipotizzare la *creazione continua* di materia nell'universo; si dimostra che in questo modo diventa non piú necessario supporre condizioni iniziali particolari per escludere la possibilità di modelli del tipo di quello di Gödel e dedurre quindi la validità del principio di Mach, ma, in virtú appunto di tale creazione, si ha una progressiva eliminazione di anisotropie e inomogeneità e, date condizioni iniziali qualsiasi, l'universo evolve verso uno stato di sempre minore vorticità, ovvero di sempre maggiore validità del principio di Mach. Si tratta naturalmente di una soluzione provvista di un alto grado di arbitrarietà, nella misura in cui non è ancora stato dato di osservare (e difficilmente potrebbe comunque venir osservato) alcun fenomeno di creazione continua; valga tuttavia a dare un'idea del livello di astrattezza e complessità che caratterizza le ricerche in questo settore.

7. *L'azione a distanza e la radiazione.*

Risale a Carl Friedrich Gauss la proposta di riformulare la teoria dell'azione a distanza di Newton (come si è visto, questa locuzione va presa con molta cautela: il fatto è che venne intesa come teoria dell'azione a distanza malgrado Newton), azione che «si propagava» in modo istantaneo, per mezzo di una teoria nella quale l'azione si propaghi a distanza, ma ad una velocità alta ma finita: quella della luce nel vuoto. Il suggerimento di Gauss fu ripreso con interessanti sviluppi da Schwarzschild [1903*a* e *b*], Tetrode [1922] e Fokker [1929], che mostrarono come effettivamente la teoria del campo elettromagnetico potesse essere costruita in accordo con tale suggerimento, in maniera matematicamente consistente, coerente con le equazioni di Maxwell e in accordo con l'esperienza per quanto riguarda i fenomeni stazionari e quasi stazionari. Le difficoltà cominciavano quando si trattavano le cosiddette correnti rapidamente variabili, e quindi i fenomeni di emissione e assorbimento di radiazione elettromagnetica.

La formulazione di Schwarzschild e Fokker ha il pregio di partire da un principio variazionale di minima azione, dal quale si può ricavare il potenziale elettromagnetico $A_\mu^{(x)}(Y)$ generato da una carica e_x posta in x, in un generico punto Y:

$$A_\mu^{(x)}(Y) = e_x \int \delta(s_{XY}^2) g_{\mu\rho}\, dX^\rho$$

essendo $g_{\mu\rho}$ il tensore metrico, X^ρ le coordinate di X e s_{XY}^2 il quadrato della distanza (in quattro dimensioni) tra X, generico punto della linea d'universo di x, e Y. Semplicemente da questo principio le equazioni di Maxwell seguono come conseguenza, ma con in piú una prescrizione precisa sulla scelta di soluzioni avanzate o ritardate, scelta che, come si è visto, non è invece determinata dalle sole equazioni di Maxwell; tale prescrizione è quella appunto già accennata, che consiste nel prendere la media aritmetica (2*c*) delle solu-

zioni (2*a*) e (2*b*). Qual è ora la difficoltà? L'esperienza mostra che una carica elettrica in moto accelerato è soggetta a una forza di frenamento simultanea all'accelerazione, e la teoria dell'azione a distanza prevede che una carica elettrica in moto accelerato senza che siano presenti altre cariche nello spazio, non sia soggetta ad alcuna forza. Anche in presenza di altre cariche i meccanismi connessi con la soluzione (2*c*) non sono sufficienti a spiegare la forza in questione, detta appunto *reazione di radiazione* (è infatti il meccanismo che permette l'emissione di radiazione elettromagnetica da parte di una carica accelerata).

Wheeler e Feynman [1945; 1949], che si occuparono del problema, ripresero un precedente suggerimento di Tetrode [1922] e di Lewis [1926] e proposero l'ipotesi che l'atto di irradiare fosse strettamente connesso con la presenza di un potenziale assorbitore, cioè di qualcosa in grado di assorbire la radiazione irradiata. Una delle loro tesi fondamentali è infatti che la forza di reazione radiativa su una carica accelerata (sorgente) sia dovuta all'azione su di essa dei campi avanzati relativi alle particelle dell'assorbitore; necessaria a questa ipotesi è l'esistenza nell'universo di un numero sufficientemente alto di assorbitori, cosí da poter applicare a questi alcune tecniche di calcolo proprie della meccanica statistica. Se ciò accade, cioè se l'intero universo è un assorbitore perfetto (ipotesi che viene formalmente tradotta assumendo che sia, ovunque nell'universo,

$$\sum_{x} (F_{(x)}^{(\mathrm{rit})} - F_{(x)}^{(\mathrm{av})}) = 0$$

dove la somma è estesa a tutte le cariche in esso presenti) si può allora dimostrare che il campo che agisce su x può essere cosí scritto:

$$\sum_{x' \neq x} F_{(x')}^{(\mathrm{rit})} + \frac{1}{2} (F_{(x)}^{(\mathrm{rit})} - F_{(x)}^{(\mathrm{av})});$$

il primo termine consiste nella somma di tutti i contributi ritardati (realmente osservati) provenienti da tutti i corpi diversi da x; è il campo esterno che agisce su x per effetto della presenza delle altre particelle e non è influenzato dal moto istantaneo di x. È invece il secondo termine quello che proviene dal moto di x, ma come risultato della risposta di tutti gli assorbitori esistenti nell'universo: ciò rende conto della reazione di radiazione e conduce a equazioni del moto per una particella carica, in un mezzo completamente assorbente, del tutto equivalenti alle usuali equazioni dell'elettrodinamica. Wheeler e Feynman, pur rimanendo spesso ancora su un terreno piuttosto qualitativo, vanno naturalmente molto oltre questi pochi cenni e si pongono problemi piú specificamente riguardanti la natura della radiazione che eccedono i limiti e gli scopi che qui ci si propongono. Ci sembra tuttavia utile l'avervi accennato perché si tratta di un esempio tipico di teoria fisica consistente ma non ancora soddisfacentemente sviluppata.

Qual è l'atteggiamento dei fisici contemporanei rispetto al dilemma azione a distanza / campo? È opinione di chi scrive che i lavori di Wheeler e Feynman

e i successivi di Hoyle e Narlikar non siano affatto popolari tra i fisici, fra i quali prevale un atteggiamento nettamente piú agnostico e/o conformista.

In primo luogo va osservato che la teoria dei campi ha, in qualche modo, prevalso. Sintomatico è quanto scrive Weyl all'inizio del lavoro nel quale per la prima volta tenta un'unificazione dell'elettromagnetismo con la teoria della gravitazione; nel definire, seguendo Riemann, un continuum metrico come una varietà i cui punti siano rappresentabili con continuità mediante tre parametri x_1, x_2, x_3 e nella quale valga il teorema di Pitagora, egli cosí giustifica l'ipotesi di far valere quest'ultimo soltanto nella sua forma infinitesima: «Del tutto in accordo con lo spirito della moderna fisica dell'azione da vicino (*Nahewirkungsphysik*) noi supponiamo che il teorema di Pitagora sia strettamente valido soltanto nell'infinitamente piccolo» [1918, p. 465]. Del resto, decisivo era stato in proposito il ruolo della relatività: praticamente da sempre vi era alla base della fisica l'idea che la situazione di un sistema ad un istante determinasse quella dell'istante successivo (con questa espressione s'intende ovviamente la situazione dopo un tempo arbitrariamente piccolo); in particolare che l'interazione ad un istante avesse il suo effetto all'istante successivo; dal momento che uno dei nuovi dettami concettuali cruciali della relatività (speciale, 1905) fu quello di considerare spazio e tempo sullo stesso piede, di trattarli cioè ovunque in maniera equivalente, il gioco era fatto: nulla sembrava piú naturale del fatto che la situazione in un punto dello spazio-tempo determinasse quella nei punti di spazio-tempo vicini, e per descrivere un tale comportamento, come si è visto, le equazioni di campo sono lo strumento naturale.

Cosí, l'etere scompariva sempre piú definitivamente dalla coscienza della comunità scientifica, non appena il programma di ricerca iniziato dai lavori di Einstein diventava prevalente sul programma di Lorentz; anche se, come nota Zahar [1973], il programma di ricerca di Lorentz, che comprendeva una teoria elettromagnetica della materia consistente con le esperienze di Michelson e Morley e con altre successive, rimase progressivo fino a dopo il 1905, finché cioè la relatività generale (1916) non portò il programma alternativo, quello di Einstein, alla previsione di fatti non altrimenti spiegati, quali lo spostamento del perielio di Mercurio. Su questa strada in verità si potrebbe andare anche piú lontano. Dicke [1961] ha recentemente mostrato in un divertente e istruttivo «sarebbe potuto accadere» che, «se Einstein fosse vissuto quindici anni piú tardi» la teoria di Lorentz avrebbe avuto il tempo di consolidarsi tanto da prevalere: essa avrebbe previsto correttamente la deviazione dei raggi luminosi in un campo gravitazionale, lo spostamento verso il rosso e anche la rotazione del perielio di Mercurio; il tutto mantenendo l'etere, e quindi lo spazio assoluto, l'aumento della costante dielettrica dell'etere in vicinanza delle masse, e la contrazione di Lorentz-Fitzgerald. Appassionante tema di ricerca questo per chi, in fisica, crede in maniera assoluta nelle «conferme sperimentali» delle teorie.

8. *Teoria quantistica dei campi.*

Anche altri furono tuttavia i successi che permisero alla teoria dei campi di prevalere: successi in termini di concisione, eleganza formale e chiarezza di formulazione matematica, ma anche in termini di previsioni sperimentali sorprendentemente corrette di molte branche della teoria. Ciò è vero, oltre che per l'elettromagnetismo classico in quanto tale, anche per il successivo massiccio sviluppo della teoria quantistica dei campi. Infatti fin dagli anni della creazione della meccanica quantistica (il terzo decennio di questo secolo) ci si è sforzati di edificare, parallelamente e coerentemente, una teoria quantistica dei sistemi comprendenti non solo particelle, ma anche campi, e in primo luogo il campo elettromagnetico. Scrivono Heisenberg e Pauli in uno dei lavori che fondano (dopo i primi contributi di Dirac) la teoria quantistica dei campi: «Non è stato sinora possibile nella teoria quantistica connettere in maniera priva di contraddizioni le leggi della meccanica e dell'elettrodinamica e le interazioni elettrostatiche e magnetostatiche da una parte con le interazioni mediate dalla radiazione dall'altra, cosí da trattarle da un punto di vista unitario. In special modo non si è riusciti a trattare in maniera corretta la velocità di propagazione finita delle azioni di forza elettromagnetiche. Scopo del presente lavoro è quello di colmare queste lacune. Per raggiungerlo sarà necessario approntare un formalismo relativisticamente invariante che permetta di trattare l'interazione tra materia e campo elettromagnetico, cosí come quella tra materia e materia» [1929, p. 1]. Cominciava cosí la costruzione di una delle parti piú ampie e articolate della moderna fisica teorica, quella che va sotto il nome di teoria quantistica dei campi. In essa il concetto di campo ricopre, com'è del tutto evidente, il ruolo fondamentale: il campo stesso, oggetto fisico cui è associato un numero infinito di gradi di libertà, e viene pertanto assimilato a un sistema di infiniti oscillatori, viene sottoposto al processo di quantizzazione: le funzioni che lo rappresentano diventano operatori (o, piú precisamente, distribuzioni a valori operatoriali) in un opportuno spazio di Hilbert (cosí come nella meccanica quantistica *strictu sensu* le osservabili – posizione, momento, ecc. – diventano operatori) e i quanti del campo sono le particelle ad esso associate: i *fotoni*, o quanti di luce, appaiono cosí come i quanti del campo elettromagnetico, ovvero quelle entità discrete che si portano dietro la quantità minima possibile, per ogni frequenza, di energia del campo. «Se i fotoni emergono in maniera cosí naturale dalla quantizzazione del campo maxwelliano, si è portati a chiedersi se altre particelle, la cui esistenza viene osservata in natura, sono anch'esse connesse a campi di forza da un analogo procedimento di quantizzazione. Su questa base Yukawa predisse l'esistenza dei mesoni π dalla conoscenza dell'esistenza delle forze nucleari» scrivono Bjorken e Drell [1965, p. 2] nell'introduzione al loro testo di teoria relativistica dei campi quantizzati. Cosí, infatti, negli anni '30 i mesoni π (indicati con i simboli π^+, π^0, π^- a seconda della loro carica elettrica; vennero in seguito scoperti altri tipi di mesoni, associati ad altri campi, ecc.) vengono introdotti come i quanti

del campo che viene associato alle interazioni nucleari, quelle cioè che tengono insieme le varie particelle che convivono nel nucleo di quasi ogni atomo (protoni e neutroni, denominati complessivamente *nucleoni*); tale campo, la cui introduzione è appunto legata al nome di Yukawa, ha un *range* dell'ordine delle dimensioni del nucleo – circa 10^{-13} cm – e dà luogo a un'interazione tra nucleoni e mesoni assai piú forte di quella elettromagnetica (legata al fatto che alcune di tali particelle posseggono anche carica elettrica non nulla). Qui è l'origine delle prime difficoltà che si riveleranno presto molto rilevanti; difficoltà legate essenzialmente all'apparire, nel formalismo matematico che esprime la teoria, di quantità infinite, le cosiddette *divergenze* – tipicamente integrali e serie che non convergono – la cui presenza fa sí che molti risultati perdano significato. L'insieme di procedure e prescrizioni escogitate per eliminare le divergenze (procedure di rinormalizzazione) è riuscito solo parzialmente a dar luogo a un quadro soddisfacente; esistono oggi varie formulazioni della teoria quantistica dei campi, a vari livelli di articolazione e di rigore, nessuna delle quali riesce comunque a inquadrare un numero sufficiente di dati sperimentali in uno schema esplicativo unitario. Esistono anzi addirittura difficoltà molto grosse a livello teorico: non è possibile a tutt'oggi neppure affermare con sicurezza l'esistenza di un modello, in dimensioni realistiche, cioè tre spaziali e una temporale, di teoria di campo rigorosa che descriva un'interazione non banale (che non descriva cioè semplicemente l'esistenza separata di particelle libere).

Assai piú promettenti erano stati i successi dell'elettrodinamica quantistica, cioè della quantizzazione del campo elettromagnetico (per il quale la costante di accoppiamento, che misura in modo preciso la forza dell'interazione, è assai minore); tipicamente la previsione dello spostamento dei livelli energetici di un elettrone di un atomo idrogenoide (*Lamb shift*) fu sperimentalmente confermata con grande precisione.

Si comprende a questo punto come lo sforzo di chi ha tentato di costruire parallelamente una teoria delle interazioni basata sul concetto di azione a distanza si sia per il momento limitato al caso delle interazioni elettromagnetiche, sforzandosi di ottenere una formulazione in grado di riprodurre tutti i risultati significativi e sperimentalmente corroborati ottenuti dalla formulazione campistica.

A questo punto la distinzione fra le due teorie dipende in maniera cruciale dal significato che si annette a questa parola. È immaginabile un esperimento che distingua fra la teoria di Wheeler e Feynman e quella campistica usuale? In linea di principio l'esperimento esiste: occorrerebbe esaminare il comportamento di un dipolo oscillante in un universo non completamente assorbente, meglio ancora in un universo dove solo il dipolo esista; secondo la formulazione usuale il dipolo dovrebbe emettere radiazione esattamente come fa nel nostro universo, secondo la teoria di Wheeler e Feynman no, e anzi in un universo altrimenti vuoto non dovrebbe emettere assolutamente nulla perché l'emissione di radiazioni dipende in tale teoria dalla presenza di assorbitori. Ma è a tutti evidente quanto sia arduo annettere valore discriminante a un

tale esperimento. Tuttavia le due teorie sono concettualmente radicalmente diverse, non solo per ciò che riguarda la loro impostazione, ma ancor piú se si considerano gli ulteriori sviluppi elaborati, sulla base delle idee di Wheeler e Feynman, da Hoyle a Narlikar [1969]. Differenze cruciali appaiono infatti nella stessa quantizzazione del campo, che, nella formulazione di questi ultimi autori, non può neppure venir effettuata dal momento che il campo è completamente determinato dalle particelle presenti, non ha cioè gradi di libertà propri; Hoyle e Narlikar riescono tuttavia a ottenere esattamente almeno alcuni di quei risultati cui si perviene usualmente mediante appunto la quantizzazione del campo, anche se ciò al prezzo di supporre opportuni modelli di universo.

Ciò è, a parere di chi scrive, assai istruttivo per la storia delle teorie fisiche e della loro pretesa falsificabilità. È facile previsione che dalla teoria dell'azione a distanza si possano ricavare, se si dispone di sufficiente abilità, con l'aggiunta eventuale di qualche ipotesi ad hoc, tutte le previsioni osservative corrette proprie anche della teoria dei campi. Il giudizio su quale delle due teorie contenga piú ipotesi ad hoc, quale sia la piú naturale, quale la piú utile, con tutta l'ambiguità che è ormai connessa a questi termini, appare sempre piú lontano e sfocato. [A. S.].

Bjorken, J. D., e Drell, S. D.
1965 *Relativistic Quantum Fields*, McGraw-Hill, New York.

Dicke, R. H.
[1961] *Mach's Principle and Equivalence*, in *Rendiconti del XX Corso della Scuola Internazionale di Fisica «Enrico Fermi»*, Academic Press, New York 1962.

Dijksterhuis, E. J.
1950 *De Mechanisering van het Wereldbeeld*, Meulenhoff, Amsterdam (trad. it. Feltrinelli, Milano 1971).

Fokker, A. D.
1929 *Ein Invarianter Variationssatz für die Bewegung mehrerer elektrischen Massenteilchen*, in «Zeitschrift für Physik», LVIII.

Gödel, K.
1949 *An example of a new type of cosmological solutions of Einstein's field equations of gravitation*, in «Reviews of Modern Physics», XXI.

Hall, A. R.
1954 *The Scientific Revolution 1500-1800. The Formation of the Modern Scientific Attitude*, Longmans, London (trad. it. Feltrinelli, Milano 1976).

Heisenberg, W., e Pauli, W.
1929 *Zur Quantendynamik der Wellenfelder*, in «Zeitschrift für Physik», LIX.

Hesse, M. B.
1961 *Forces and Fields. The Concept of Action at a Distance in the History of Physics*, Nelson, London-New York (trad. it. Feltrinelli, Milano 1974).

Hoyle, F., e Narlikar, J. V.
1963 *Mach's principle and the creation of matter*, in «Proceedings of the Royal Society. Series A: Mathematical and Physical Sciences», CCLXXIII.
1969 *Electrodynamics of direct interparticle action. I. The quantum mechanical response of the universe*, in «Annals of Physics», LIV.

Jammer, M.
1957 *Concepts of Force. A Study in the Foundations of Dynamics*, Harvard University Press, Cambridge Mass. (trad. it. Feltrinelli, Milano 1971).

Kelvin (W. Thomson, Lord)
[1884] *Baltimore Lectures on Molecular Dynamics and the Wave Theory of Light*, Clay, London 1904.

Lewis, G. N.
1926 *The nature of light*, in «Proceedings of the National Academy of Sciences», XII.

Mach, E.
1883 *Die Mechanik in ihrer Entwickelung historisch-kritisch dargestellt*, Brockhaus, Leipzig (trad. it. Boringhieri, Torino 1968).

Maxwell, J. C.
1861 *On physical lines of force*, in «Philosophical Magazine», XXI.
[1864] *A dynamical theory of the electromagnetic field*, in «Royal Society. Philosophical Transactions», CLV (1865).
1873 *Treatise on Electricity and Magnetism*, Clarendon Press, Oxford (trad. it. Utet, Torino 1973).

Newton, I.
1704 *Opticks: or, a Treatise of the Reflexions, Refractions, Inflexions and Colours of Light*, Smith and Walford, London; ed. Innys, London 1730.
1713 *Philosophiae Naturalis Principia Mathematica*, Crownfield, Cambridge 1713² (trad. it. Utet, Torino 1965).

Schwarzschild, F.
1903*a* *Zur Elektrodynamik. I. Zwei Formen des Princips der kleinsten Action in der Elektronentheorie*, in «Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse».
1903*b* *Zur Elektrodynamik. II. Die Elementare elektrodynamische Kraft*, *ibid.*

Tetrode, H.
1922 *Über den Wirkungzusammenhang der Welt. Eine Erweiterung der klassischen Dynamik*, in «Zeitschrift für Physik», X.

Thom, R.
1972 *Stabilité structurelle et morphogénèse. Essai d'une théorie générale des modèles*, Benjamin, Reading Mass.

Weyl, H.
1918 *Gravitation und Elektrizität*, in «Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin».

Wheeler, J. A., e Feynman, R. Ph.
1945 *Interaction with the absorber as the mechanism of radiation*, in «Reviews of Modern Physics», XVII.
1949 *Classical electrodynamics in terms of direct interparticle action*, in «Reviews of Modern Physics», XXI.

Zahar, E.
1973 *Why did Einstein's Programs supersede Lorentz's?*, in «British Journal for the Philosophy of Science», XXIV.

Analogamente ad altre nozioni, per esempio **energia**, quella di forza si è progressivamente svincolata da alcuni dei riferimenti intuitivi – legati ad esperienze (cfr. **dato**, **empiria/esperienza**) materiali umane o naturali – ed è divenuta, grazie a un procedimento di rigorizzazione tipico della **conoscenza** scientifica, un potente **concetto** unificatore nella **fisica**. Entro il **paradigma** newtoniano ha rappresentato, per la meccanica (cfr. **moto**), il correlato della nozione di causa (cfr. **causa/effetto**, **determinato/indeterminato**) ove la concezione dell'azione a distanza, centrale nello studio

della **gravitazione**, si è imposta grazie al tipo di **calcolo** che essa consente e al conseguente potere predittivo (cfr. **anticipazione, previsione**). Ma, anche in connessione con l'indagine sistematica di altri fenomeni (cfr. **fenomeno**) – elettrici, magnetici, ecc. – è emersa la concezione di campo, ora alternativa, ora complementare. La coppia forza/campo ha quindi svolto un diverso ruolo nella **ricerca** scientifica a seconda delle particolari opzioni filosofiche e metodologiche (cfr. **coppie filosofiche, filosofia/filosofie, metodo**). Corrispettivo fisico del passaggio dal globale al locale (cfr. **locale/globale**), quello dall'azione a distanza alla concezione campistica interessa sia la **relatività** sia la fisica dei **quanti** e ha conseguenze nelle varie **cosmologie** e nelle stesse concezioni generali del **mondo**. La complessa **dialettica** che inerisce a tale coppia non è piú riducibile a un qualche **esperimento** cruciale che permetta di decidere per l'una o l'altra concezione (cfr. **verificabilità/falsificabilità**), ma s'iscrive in una visione piú articolata della **spiegazione** scientifica e delle procedure di controllo delle teorie (cfr. **teoria/modello**).

Moto

La definizione di moto quale è data dai comuni dizionari può essere accettata anche nel linguaggio scientifico: cosí, per lo Zingarelli si parla di un «moto di un corpo, allorché questo occupa successivamente diverse posizioni nello spazio»; per il *Vocabolario della lingua italiana* di Giacomo Devoto e Gian Carlo Oli «il moto è la variazione continua della posizione di un corpo, o di una parte di esso, rispetto ad un altro». In sostanza, ciò che affermano innanzi tutto le predette definizioni è che ogni fenomeno di moto si svolge nello spazio e nel tempo: ora, proprio l'evoluzione di questi concetti primari di spazio e di tempo, prodottasi dalle epoche piú remote ad oggi, è causa della differenziazione che detti concetti, e quindi quello di moto, possono presentare nel linguaggio scientifico e in quello comune.

Da entrambe le definizioni di moto sopra ricordate, poi, appare subito che la nozione di moto, cosí come quella di quiete, è per sua natura *relativa*, in quanto parlare della posizione di un corpo C ha senso solo se viene definito il sistema S rispetto al quale il corpo C è riferito, ossia se si definisce il cosiddetto «sistema di riferimento»: questo può essere costituito da una terna di assi cartesiani ortogonali $(O; x, y, z)$ solidali all'osservatore che considera il moto. Se P è un punto qualsiasi del corpo mobile C, le coordinate di P rispetto alla terna S varieranno nel tempo e si potrà scrivere:

$$(1) \qquad x=x(t); \qquad y=y(t); \qquad z=z(t)$$

o compendiosamente $P=P(t)$ se si indica con t la variabile temporale quale è misurata dall'osservatore solidale con S. Al variare di t variano le posizioni delle proiezioni dei tre punti P_x, P_y, P_z, proiezioni di P sui tre assi, ed il moto di P si dice composto dei moti di P_x, P_y, P_z. La rapidità con la quale varia la posizione di P sulla sua traiettoria, ossia sulla linea di equazioni (1), è definita dal vettore

$$(2) \qquad \lim_{\Delta t\to 0}\frac{P(t+\Delta t)-P(t)}{\Delta t}=\frac{dP}{dt}=\mathbf{v}$$

le cui componenti secondo i tre assi x, y, z sono dx/dt, dy/dt, dz/dt di guisa che se $\mathbf{i}_x$, $\mathbf{i}_y$, $\mathbf{i}_z$ sono i versori degli assi stessi, è

$$(2') \qquad \mathbf{v}=\frac{dx}{dt}\mathbf{i}_x+\frac{dy}{dt}\mathbf{i}_y+\frac{dz}{dt}\mathbf{i}_z,$$

o anche, se ds è la lunghezza dell'arco elementare della traiettoria di P descritto da P nel tempuscolo dt, e $\boldsymbol{\tau}$ il versore della tangente in $P(t)$ a detta linea,

$$(2'') \qquad \mathbf{v}=\frac{ds}{dt}\boldsymbol{\tau}.$$

Il vettore $\mathbf{v}$ ora definito è la velocità di P, e la sua definizione corrisponde alla nozione intuitiva di rapidità del moto, mentre dalla (2′) appare che la velocità

del punto è data in ogni istante dalla risultante delle velocità che in quell'istante competono ai moti componenti.

Ora, è facile rendersi conto che, se si considera il moto dello stesso punto P rispetto a un altro sistema di riferimento S' $(O'; x', y', z')$, la cui posizione cambia rispetto a S (e quindi è in moto relativamente a S), i moti di P rispetto a S e a S' sono diversi, in quanto sono diverse le traiettorie e diverse le velocità corrispondenti: cosí, se si considera un treno che si muove in curva, e si suppone che gli assi delle due rotaie siano in uno stesso piano orizzontale e disposti secondo due circonferenze concentriche, per un osservatore, seduto nel suo scompartimento mentre un altro viaggiatore passeggia lungo il corridoio, la traiettoria del baricentro di questo secondo viaggiatore può apparire rettilinea; un osservatore esterno al treno e immobile (rispetto alla terra) vedrà invece circolare la traiettoria stessa.

Questa non-invarianza del moto di un punto, al variare del riferimento, portò il filosofo Parmenide d'Elea, il filosofo «venerando e insieme terribile» nella sua grandezza, come disse Platone [*Teeteto*, 183e], a negare il movimento stesso, che sarebbe pertanto solo un'apparenza dei sensi: ventiquattro secoli piú tardi, però, Einstein dimostrò che se non è invariante la linea che rappresenta la traiettoria nel riferimento spaziale (che costituisce un continuo tridimensionale nelle coordinate x, y, z), è invariante la linea che la rappresenta nel riferimento spazio-tempo (che è un continuo quadridimensionale nelle coordinate x, y, z e t), e pertanto il concetto di moto, anche nella concezione di Parmenide, può acquisire diritto di esistenza. È naturalmente ovvio domandarsi se, noto il movimento di un punto P rispetto ad un riferimento S, e noto il moto del sistema S rispetto a un altro riferimento S', si possa determinare il moto di P relativamente a S'; è questo un problema fondamentale della scienza che studia il moto, la meccanica, e la risposta data al quesito sin dai tempi piú antichi è senz'altro sí, ma il risultato di questa composizione dei moti di P in S e di S in S' risente dell'evoluzione del concetto di moto, cui si è accennato.

Il problema dei problemi però, la cui soluzione è stata l'aspirazione per molti millenni, e tale continua ad essere oggi, perché una soluzione generale, rigorosa, non esiste tuttora, è quello della spiegazione del moto, ossia della determinazione della causa che produce il movimento, e degli effetti che conseguono a date cause, e pertanto delle leggi del moto, il che comporta la spiegazione dei moti che si producono in natura (moto dei corpi celesti, che da sempre hanno attirato l'attenzione degli uomini; locomozione animale) e, inversamente, delle cause che debbono essere generate per ottenere determinati movimenti. Questo ha richiesto l'introduzione di due altri concetti primitivi, in aggiunta a quelli di spazio e di tempo, e precisamente dei concetti di forza e di massa. Anche questi hanno avuto, particolarmente in epoca recente, un'evoluzione profonda; è opportuno però subito precisare che le evoluzioni, alle quali finora si è accennato, hanno portato sí a una piú corretta descrizione dei fenomeni, e a una loro rappresentazione per mezzo di modelli matematici piú comprensivi, che sono cioè applicabili a categorie sempre piú vaste di fenomeni, ma non ad una vera e propria loro comprensione. Vale a questo riguardo quanto dice Werner Heisenberg [1932]:

«Quanto piú vasto è il campo che la fisica, la chimica e l'astronomia ci rivelano, tanto piú noi sogliamo sostituire l'espressione "spiegazione" della natura con la piú modesta espressione "descrizione" della natura, e tanto piú chiaro diventa che questo progresso non è dato da un sapere diretto, ma da una comprensione analitica. Con ogni grande scoperta, – e lo si può seguire specialmente nella fisica moderna, – diventano minori le pretese degli scienziati a una comprensione del mondo nel senso originario» (trad. it. p. 11).

In quanto segue non saranno considerati i vari problemi nei quali si scinde il problema generale, quale sopra è stato specificato, perché questo richiederebbe oltre che uno spazio molto maggiore di quello che può essere concesso a un articolo d'enciclopedia, anche competenze specifiche diverse, ed in effetti quanto qui è omesso potrà essere ritrovato in altri articoli di questa stessa *Enciclopedia*; ci si limiterà soltanto ad indicare quanto è necessario per una prima visione generale del problema, e questo sarà fatto in due paragrafi, nel primo dei quali sarà esposto il modo in cui si è andato evolvendo nel tempo il concetto di moto, ed indicate le cause che hanno condotto ad una tale evoluzione; nel secondo saranno prese in considerazione le modalità con le quali si realizzano i moti naturali, limitatamente alla locomozione acquatica degli esseri viventi. Questa ovviamente soddisfa alle leggi fisiche, ma le modalità con le quali essa potrebbe essere realizzata non sono univocamente determinate dall'ambiente in cui la specie considerata vive, mentre è enorme il numero delle specie e sottospecie che costituiscono il mondo animale, ed enorme è il rapporto delle masse degli organismi viventi (da 10^{-13} g per i Micoplasmi a 10^{8} g per le Balene blu, corrispondente quindi ad un rapporto delle masse estreme uguale a 10^{21}). È generalmente ammesso che queste modalità corrispondano a condizioni ottime per la locomozione soggetta a determinate condizioni, dipendenti o dall'ambiente o dalle dimensioni dell'organismo, o da altre circostanze, ed in effetti nei casi in cui si riesce a determinare l'efficienza dei mezzi propulsivi si constata che tale efficienza è la piú grande, o prossima alla piú grande (nei limiti d'approssimazione del calcolo), che potrebbe aversi. Poiché si ritiene che la vita animale sia comparsa sulla Terra da circa 10^{9} anni, questa ottimizzazione è generalmente attribuita a quella che si suole chiamare selezione naturale: la natura agirebbe come un gigantesco calcolatore, dotato di memoria prodigiosa, che procede con un programma, che applica il metodo delle successive approssimazioni per determinare il sistema ottimo per ogni animale, od ogni specie di animali; ma anche questa frase può descrivere il fenomeno, non spiegarlo.

1. *Il moto come è concepito e studiato dall'uomo. Evoluzione dei concetti meccanici.*

1.1. Dall'evo antico a Galileo.

Non s'intende qui fare la storia del modo col quale nell'uomo primitivo si formarono i concetti di spazio e di tempo, che probabilmente furono conseguenza delle sensazioni e delle elementari reazioni fisiche da lui provate per effetto

del corso quotidiano del sole, di ciò che vedeva quando variava la sua posizione portandosi da un luogo all'altro, della fatica che doveva fare quando compiva tali spostamenti, ecc... È del pari verosimile che i fenomeni che piú attrassero la sua attenzione e piú lo impressionarono furono i fenomeni celesti, in particolare il moto degli astri, ed in effetti lo studio di tali moti fu tra i primi ad essere coltivato: già cinquemila anni or sono era fiorente l'astronomia caldaica, alla quale diedero efficaci contributi calcolatori e osservatori abilissimi, rimanendo peraltro allo stato puramente descrittivo. L'astronomia acquistò poi dignità di scienza sia ad opera dei Greci, e in particolare dei pitagorici, che, brillanti cultori ed amanti della geometria, la collegarono con questa, sia degli Egiziani: il trattato nel quale sono raccolti gli studi astronomici piú importanti fondati sulla concezione dell'immobilità della Terra è la famosa μαθηματιχὴ σύνταξις, detta poi *Almagesto* (dall'articolo arabo *al* 'la', e dalla parola greca μεγίστη 'la piú grande') ed è opera di Claudio Tolomeo. Da esso risulta un sistema solare (nella costituzione quale era nota a quei tempi) che soddisfa alle regole della filosofia e della fisica aristoteliche: secondo dette regole, ogni corpo ha il posto che la natura gli ha assegnato, e qualora si muova esso raggiunge via via posti ben determinati, e precisamente i corpi materiali terrestri debbono avere un moto *naturale* verticale, mentre per i corpi celesti il moto naturale è quello circolare uniforme.

Nella concezione originaria di Eudosso di Cnido ogni corpo celeste è pensato incastonato in una sfera girevole su due poli imperniati sopra un'altra sfera, e cosí via; le sfere ruotano e trascinano nel loro moto rotatorio gli astri in traiettorie circolari. La Terra è nel centro comune di tutte le varie sfere, mentre i moti dei pianeti si ottengono dalla composizione di moti circolari uniformi; le sfere concentriche allora ammesse erano ventisette, una essendo quella delle stelle fisse. Questa concezione fu pienamente accolta da Aristotele il quale, per giustificare che il moto degli astri esterni non veniva a comunicarsi a quelli interni, aggiunse altre sfere dotate di moti inversi, dette «reagenti», portando cosí il numero complessivo delle sfere a cinquantacinque. Per Tolomeo le sfere celesti erano pure figure geometriche, mentre per Aristotele le sfere erano materiali, formate da una sostanza semplice, unica e diversa dagli elementi costituenti il mondo sublunare, chiamata «etere», e dai posteri la *quinta essentia*. Non è qui il caso di inoltrarsi nel groviglio di tutta la complessa costituzione dell'universo in dette rappresentazioni [per questo si veda Neugebauer 1975], resa cosí aggrovigliata non tanto dalla scelta non appropriata del centro del sistema, quanto dalla necessità di *salvare le apparenze*, ossia di ottenere una buona corrispondenza tra i risultati deducibili da detta costituzione e i risultati delle osservazioni astronomiche, compreso il fenomeno della precessione degli equinozi scoperto da Ipparco nel II secolo a. C. dal confronto delle sue determinazioni di longitudine di stelle fisse con analoghe osservazioni fatte centocinquant'anni prima da Aristillo e Timocari. Quello su cui importa porre l'attenzione sono i concetti fondamentali quali risultano da tutto il complesso di ricerche sopra accennato, ed essenzialmente quali appaiono dalla filosofia e dalla fisica aristoteliche, e che ebbero un influsso determinante sul pensiero scientifico fino a Galileo e oltre.

Si osservi anzitutto che detta filosofia postula l'esistenza di un punto naturale, e perciò privilegiato, di osservazione: il centro della Terra, situato nel centro dell'Universo; e di conseguenza anche l'esistenza di un sistema privilegiato di riferimento, solidale con la Terra stessa, e perciò di un moto assoluto, definito come il moto quale appare rispetto a detto riferimento. È opportuno subito dire a questo riguardo che, se il geocentrismo cadde con Copernico e Galileo, il principio di esistenza di un sistema di riferimento privilegiato, e quindi del moto assoluto, continuò ad essere ammesso anche in epoche molto posteriori, e precisamente fino al secolo attuale; naturalmente cambiò l'ente al quale detto sistema si pensava ancorato: cosí, per Copernico il sistema privilegiato era costituito da un sistema di assi con l'origine nel centro del Sole e solidale con questo; per Newton era il sistema solidale con le cosiddette «stelle fisse», ossia quelle stelle che (almeno fino alla metà del XVIII secolo) si presentavano come punti luminosi tra i quali non si osservava da secoli alcun sensibile mutamento della loro posizione relativa; successivamente, quando le ricerche di astronomia siderale hanno messo in evidenza che nella realtà anche le cosiddette stelle fisse erano mobili, fu in un primo tempo richiamato il concetto di etere, sia pure considerato non piú come *quinta essentia* aristotelica, ma come una sostanza senza peso, senza moto, incorruttibile e inalterabile, trasparente e permeante l'Universo intero, e si pensò il sistema di riferimento per il moto assoluto solidale con detta sostanza; poi, quando il concetto di etere fu definitivamente sepolto, si ritornò alle stelle, sia pure non piú considerate come fisse, e del cui moto si cercò di tenere conto precisando il riferimento con considerazioni di media statistica.

Un altro postulato insito nella fisica aristotelica è quello dell'esistenza del tempo assoluto, per il quale la misura del tempo non dipende dallo stato di quiete o di moto dell'osservatore, e anche questo concetto del tempo assoluto continuò ad essere ammesso successivamente fino ai tempi moderni, e piú precisamente sino all'attuale secolo.

Dai due postulati ricordati discende subito il principio di composizione delle velocità, per il quale la velocità del moto composto è la somma (vettoriale) delle velocità dei moti componenti, e pertanto se S è il riferimento assoluto, S' un sistema di riferimento che ha un moto noto rispetto a S, per un punto P qualsiasi mobile sia rispetto a S sia rispetto a S', la velocità assoluta $\mathbf{v}_a$ di P, e la velocità relativa $\mathbf{v}_r$ di P (rispetto a S'), sono legate dalle

$$\mathbf{v}_a = \mathbf{v}_r + \mathbf{v}_t \tag{1}$$

essendo $\mathbf{v}_t$ la velocità di P pensato solidale con S' (velocità di trascinamento). Il principio di composizione dei movimenti secondo la (1) era ben noto agli antichi, come appare dalle ricerche degli astronomi cui sopra si è accennato. Non era invece affatto nota la *causa* del moto, od almeno le loro concezioni al riguardo erano del tutto non corrette, né d'altra parte si può dire che essi se ne siano molto preoccupati, le loro ricerche essendo essenzialmente limitate a dare una rappresentazione geometrica del moto nel tempo.

Per Aristotele ogni cosa esistente, terrestre o celeste, ha un suo luogo naturale ed è composta di materia, forma e movimento, ovvero di una o di due di tali

proprietà; Dio è pura forma, perciò è immateriale, immobile e incorruttibile; il mondo sublunare ha invece tutti e tre gli attributi, perciò è corruttibile in quanto materia e il suo moto ha un termine; i corpi celesti godono solo delle proprietà di forma e di movimento, e conservando perciò, come attributo divino, la immutabilità, il loro moto è eterno. Ne consegue che il moto rettilineo (che ha un termine), quale si osserva nella caduta dei gravi verso il centro della Terra, o nell'ascesa della fiamma verso l'ultima sfera sublunare, mancando della caratteristica dell'eternità, è proprio della materia corruttibile; il moto circolare invece non ha termine ed è quindi l'unico che possa competere ai corpi celesti.

D'altra parte il concetto di forza era pure ben noto agli antichi, e in particolare ad Aristotele, anche se detto concetto connesso con la parola 'forza' (δύναμις, ἰσχύς) non corrisponde perfettamente al significato che ha assunto questa parola nella meccanica moderna: la differenza consiste essenzialmente nel fatto che da Aristotele le forze sono considerate unicamente sotto l'aspetto di grandezze scalari. «È naturale quindi che a lui, che pure aveva un concetto cosí chiaro della composizione dei movimenti (Μηχ. Προβλ., 23) e delle velocità (*ib.*, 1)», come già si è detto, «non si presentasse neppure il problema della composizione di due forze di diversa direzione, applicate a uno stesso punto» [Vailati 1897, p. 945, nota 1], e questo, forse, per lo stesso processo col quale la nozione di forza si è imposta all'uomo.

Al pari dei concetti di spazio, di tempo, di moto, in effetti anche quello di forza è derivato da sensazioni fisiche, e precisamente dalla sensazione muscolare di sforzo, che occorre esercitare, ad esempio, per sollevare un corpo. Il processo logico col quale dalla sensazione fisica di sforzo si risale al concetto astratto piú generale di forza è abbastanza semplice: poiché l'effetto piú appariscente, che si osserva quando si applica con le mani uno sforzo a un corpo, è quello di metterlo in moto, cosí si è portati naturalmente a chiamare forza ogni circostanza capace di produrre gli stessi effetti meccanici, ossia le medesime conseguenze sul moto o sulla quiete, rispetto a un riferimento qualsiasi, del corpo cui è applicata, che vi sarebbero prodotti da un conveniente sforzo muscolare. Si riconosce facilmente, in base a questo criterio, che l'azione che una corrente fluida esercita sopra gli ostacoli che essa incontra è una forza: cosí il vento fa sbattere le imposte di una finestra, ed è possibile tenerle ferme con le mani pur di esercitare un conveniente sforzo; la corrente d'acqua di un fiume trascina con sé galleggianti, ed è possibile opporsi a detta azione se dalla riva viene esercitato su di essi, con un mezzo appropriato, uno sforzo sufficiente.

Appunto per questa sua derivazione dalla sensazione di sforzo muscolare, Aristotele, e con lui tutti gli antichi, non riuscirono a fare completamente astrazione dalle proprietà che si riferiscono a questo caso speciale; cosí, allo stesso modo che, quando si parla della forza di un uomo o di un animale, si intende riferirsi al massimo sforzo che esso può esercitare, quando Aristotele accenna alla direzione secondo cui agisce una forza applicata a un determinato punto di un meccanismo, egli intende sempre riferirsi alla direzione che la forza, di cui si tratta, deve avere affinché essa riesca quanto piú possibile efficace. In effetti, quando considera forze applicate a un disco girevole intorno al suo asse, egli le

suppone sempre tangenti alle circonferenze che passano per i loro punti di applicazione, e usa la frase: «circolo secondo cui agisce, o a cui è applicata una forza» [citato *ibid.*, p. 945].

Ad ogni modo, anche per Aristotele la forza è la causa del moto, in quanto egli afferma nella *Fisica* [256a, 2] che «tutte le cose mosse risultano mosse da qualcosa», ma occorre distinguere i moti naturali, che si producono *spontaneamente*, ossia senza che ci sia contatto diretto, o indiretto per mezzo di una catena di oggetti intermedi in contatto fra loro, quali il moto dei gravi, o il moto dei corpi celesti, od anche l'ascensione delle fiamme, da quelli per i quali si ha *contiguità* tra corpo in moto (o corpo su cui si esercita la forza) e causa del moto (ossia corpo da cui proviene la forza): per i primi vale per Aristotele la spiegazione cui sopra si è accennato, per i secondi la causa è la forza secondo il concetto prima indicato. Peraltro, per ciò che riguarda lo studio delle leggi del moto, Aristotele e i suoi successori fino a Galileo sembrano essersi accontentati di grossolane descrizioni e classificazioni dei fenomeni: in effetti, tutto ciò che si legge a questo proposito nella *Meccanica* è: «Il corpo in moto si arresta, allorché la forza che lo spinge non può piú oltre agire in modo da spingerlo» [citato in Einstein e Infeld 1938, trad. it. p. 19]. Questa proposizione è stata interpretata nel senso che la causa forza ha come effetto la velocità, e pertanto tanto maggiore è la forza, tanto piú grande è la velocità che si può ottenere, ed è la velocità, di conseguenza, che indica se forze esterne agiscono o no su corpi: questa interpretazione è corrispondente alla nozione intuitiva, che può essere tratta dall'esperienza quotidiana, che un veicolo trainato da piú cavalli può essere piú veloce di quello trainato da un solo cavallo. Ora è noto che le affermazioni sopra indicate riguardo all'effetto che una forza produce sono del tutto errate; d'altra parte se la proposizione sopraddetta è interpretata nel senso che, cessando l'azione della forza che produce il moto, il moto stesso *finisce* per arrestarsi, perché sempre esistono forze che si oppongono al movimento (forze resistenti), essa è ovviamente valida, ma nulla dice riguardo alla legge del moto. Ora è indubbiamente strano che tutta la dinamica (ossia la parte della meccanica che studia il moto di un punto di un sistema in rapporto alle cause che lo determinano) aristotelica si riduca alla proposizione sopra enunciata: strano, perché il principio che sarà in epoca molto piú tarda riconosciuto come il principio fondamentale della meccanica, il principio dei lavori virtuali, era in nuce contenuto già nelle *Questioni meccaniche* piú volte citate, come dimostrato da Giovanni Vailati [1897, p. 948]. È vero peraltro che Aristotele si limita a considerare la sua applicazione a problemi di statica, a stabilire cioè le relazioni che le forze applicate debbono soddisfare perché siano in equilibrio, e pertanto non producano alcun moto del sistema cui sono applicate; ed ancora si deve osservare che egli non ha introdotto un nome speciale per designare il prodotto della forza per lo spostamento del suo punto di applicazione nella direzione cui essa è applicata, per quanto il concetto di lavoro (o di energia) appaia contenuto nei suoi ragionamenti: la parola 'energia' (ἐνέργεια) è infatti adoperata, in opposizione a δύναμις, per designare, in generale, il contrasto fra l'*esercizio* di una data facoltà e il *possesso* della facoltà stessa. Cosí, in un passo della sua opera *Del cielo* [308a, 2-3] la distinzione tra la δύναμις,

corrispondente al fatto che un grave sospeso *potrebbe* cadere, e l'ἐνέργεια, corrispondente alla sua effettiva caduta, è dimostrata per mezzo del paragone di un focolare, in cui la brace cova sotto la cenere. Si può pensare che il mancato sviluppo delle idee insite nel principio dei lavori virtuali, sia pure nella forma pensata da Aristotele, sia dovuta in parte al non sufficiente sviluppo della matematica, e in parte ai principî filosofici e religiosi che codificavano una fisica sia terrestre sia celeste, e stabilivano una cosmologia su basi apparentemente intoccabili e incrollabili. L'influenza del pensiero filosofico e religioso su quello scientifico, e viceversa, non è una caratteristica dell'epoca di Aristotele, ma sempre si è prodotta e tuttora si verifica. Forse questo è inevitabile, ma è certo pericoloso; per l'autore di questo articolo quello che importa tener presente è che il mondo fisico e il mondo spirituale sono distinti e non è lecito applicare i metodi e i concetti che valgono nell'uno per comprendere e spiegare l'altro senza una severa indagine critica: Werner Heisenberg [1948] dice: «I concetti di anima o di vita non figurano nella fisica atomica, e perciò non possono essere dedotti, nemmeno indirettamente, come conseguenze piú o meno complicate di qualche legge della natura» (trad. it. p. 153).

Ad ogni modo è certo che l'autorità di Aristotele in Europa è stata molto grande, e questo probabilmente non è stato una delle ultime ragioni per le quali errori contenuti nella sua dottrina fisica si sono perpetuati per secoli, in quanto produsse un degenere aristotelismo, caratterizzato da un rigido dogmatismo destinato a spegnere ogni tentativo di indagine sperimentale. È curioso osservare che concetti corretti contenuti nella meccanica aristotelica, e piú progrediti di quelli dai quali i moderni hanno preso le mosse per spingersi innanzi, sono stati severamente criticati e respinti, mentre altri concetti erronei sono stati piú o meno consapevolmente accettati senza critica almeno fino a Galileo. Questo è quanto accadde per il principio dei lavori virtuali: la sua affermazione che «condizione necessaria e sufficiente, perché delle grandi resistenze possano essere vinte da piccole forze, è che il meccanismo sia costruito in modo che, a un dato spostamento di quello dei suoi punti a cui la forza è applicata... corrisponda un minor cammino di quell'altro punto al cui moto si oppone l'ostacolo che si vuol superare» [citato in Vailati 1897, p. 947] fu aspramente criticata da Simone Stevin nella *Statica* (*De Beghinselen der Weegkonst*, 1586), in cui egli dichiara essere assurdo il ragionamento di Aristotele, perché questo «implica il considerare l'equilibrio come dipendente dal sussistere di condizioni che non possono effettivamente realizzarsi se non quando l'equilibrio cessa di sussistere» [cfr. Vailati 1897, p. 949]. Questa stessa argomentazione avrebbe potuto portare, invece che a respingere il ragionamento di Aristotele, a stabilire prima di Galileo e di Newton il principio d'inerzia. A questo ripudio del principio giusto si contrappone invece l'accettazione del principio che il moto naturale dei corpi celesti sia un moto circolare uniforme o deducibile dalla composizione di moti circolari. In effetti, anche nel sistema copernicano questa proprietà è rispettata: nel *De revolutionibus* [1543] Copernico dice: «Dopo ciò, ricorderemo che il movimento dei corpi celesti è circolare. Infatti la mobilità della sfera consiste nel ruotare in circolo» (trad. it. p. 47) (nella versione latina: «Post haec memorabimus corporum

caelestium motum esse circularem. Mobilitas enim sphaerae est in circulum volvi»), ed il concetto delle sfere celesti è conservato, anche se egli, a differenza di Aristotele, non le materializza, ma sostituisce al concetto di un etere o *quinta essentia* quello di «pura forma in movimento», che possiede una virtú dinamica; ed è proprio questa virtú che permette non solo di far ruotare la Terra su se stessa, ma addirittura di trascinarla nei cieli.

Con tutto questo, e per quanto i concetti di spazio, tempo e forza non presentino da Aristotele a Copernico alcuna evoluzione, neppure per quanto si riferisce alla relazione tra la causa del moto, la forza, e l'effetto prodotto, il moto stesso, il passaggio dal sistema tolemaico al sistema copernicano per la rappresentazione del sistema solare costituisce un progresso di enorme, incalcolabile importanza, anche, e forse soprattutto, perché, trasportando il centro del sistema di riferimento privilegiato dal centro della Terra al centro del Sole, ha aperto la strada a Keplero per la scoperta delle sue leggi sul moto dei pianeti, e questi, a sua volta, ha dato i fondamenti sperimentali indispensabili a Newton per la sua teoria della gravitazione universale; ossia, come ben disse Geymonat, la grandezza di Copernico non consiste tanto «nell'avere acceso una luce che brilla eterna nel cielo delle verità assolute, quanto nell'avere acceso una luce capace a sua volta di accenderne una piú grande» [1973, p. 30].

Da notare che Keplero basò i suoi calcoli principalmente sulle osservazioni astronomiche di Tycho Brahe, che fu uno dei piú accaniti oppositori di Copernico, e che una delle principali difficoltà incontrate da Keplero nella rappresentazione ed interpretazione di dette osservazioni fu dovuta proprio al fatto che egli, studiando l'orbita del centro di Marte, era pur sempre ancorato al principio aristotelico che le traiettorie dei corpi celesti fossero circonferenze. Ora, constatando che fra le posizioni di Marte calcolate e quelle osservate c'era un divario che raggiungeva gli otto primi di angolo, riconobbe che detto divario doveva essere prodotto da qualche errore, che certamente non poteva essere attribuito alle osservazioni di Brahe, perché egli stesso afferma nell'*Astronomia nova* (1609) che la divina bontà diede in Tycho Brahe un osservatore molto accurato. Fu quindi portato ad ammettere che l'orbita marziana fosse un ovale, ma poiché, dopo calcoli faticosi e lunghi, si rese conto che anche questa ipotesi non portava ad un risultato accettabile, dopo un ritorno, pure vano, all'ipotesi della traiettoria circolare, si decise per l'orbita ellittica, forse perché l'ellisse, tra le ovali, è quella piú affine al cerchio, e riconobbe che la concordanza con i risultati sperimentali diventava ottima, ammettendo che il Sole fosse in uno dei fuochi. Egli ricavò cosí quella che è nota come la prima delle tre leggi che portano il suo nome:

PRIMA LEGGE. *Le orbite dei pianeti sono ellissi, e il Sole ne occupa uno dei fuochi.*

Tale legge è da porsi accanto a quella generalmente nota come

SECONDA LEGGE. *Le aree descritte dal raggio vettore che va dal Sole a un pianeta sono proporzionali ai tempi impiegati a percorrerle.*

benché esse siano state nella realtà ottenute in ordine cronologico inverso. Alquanto posteriore è la

TERZA LEGGE. *I quadrati dei tempi impiegati dai vari pianeti a percorrere le loro orbite (durata delle rivoluzioni) sono proporzionali ai cubi dei semiassi maggiori,*

che fu pubblicata nelle *Harmonices Mundi* (1619). Benché Keplero ricavasse le sue leggi solo dalle osservazioni relative a Marte (o, se si vuole, alla Luna), egli, credendo fermamente nell'armonia del creato, era convinto che esse valessero per tutti i pianeti, ed in effetti i calcoli successivamente fatti gli diedero ragione.

Per quanto si riferisce peraltro alla causa del moto, e alla relazione tra causa ed effetto, Keplero non aveva le idee piú chiare di Copernico: anch'egli, per spiegare il moto dei pianeti, riteneva che la «virtú motrice» risiedesse nel Sole, e piú precisamente che dal Sole emanasse un flusso che, irradiandosi e colpendo i pianeti, li sospingesse nelle loro orbite. Nel *Mysterium cosmographicum* egli infatti scrive: «C'è una sola azione motrice posta nel centro di tutti gli orbi, cioè nel Sole, la quale sospinge ogni corpo, e, precisamente, con maggior forza quelli che le sono piú vicini, e piú debolmente i lontani, in proporzione dell'attenuazione della forza colla distanza» [1596, ed. 1938 p. 70].

Contemporaneo di Keplero fu Galileo Galilei, ed è Galileo che può considerarsi come l'iniziatore della creazione della scienza del moto, cioè della dinamica, intesa come scienza il cui oggetto non è la pura descrizione geometrica e temporale dei fenomeni di moto, ma l'indagine causale dei fenomeni stessi. Anche Galileo postula un tempo assoluto, ossia indipendente dal moto dell'osservatore; egli però per il primo si vale di un altro concetto fondamentale, quello di accelerazione, da lui definito come «variazione della velocità nell'unità di tempo» e pertanto rappresentata dalla

$$a=\frac{\Delta \dot{s}}{\Delta t} \tag{2}$$

se $\dot{s}(t)$ è la velocità sulla traiettoria rettilinea e $\Delta\dot{s}=\dot{s}(t+\Delta t)-\dot{s}(t)$. È da notare che Galileo non aveva a disposizione il calcolo infinitesimale, e non poteva perciò formalmente definire l'accelerazione se non nel caso in cui essa è costante nel tempo, e perciò essendo Δt un intervallo di tempo qualsiasi, il rapporto $\Delta\dot{s}/\Delta t$ non dipende da Δt; è precisamente quanto avviene nel moto di caduta dei gravi, da lui chiamato «moto uniformemente accelerato». Il concetto di accelerazione fu poi esteso da Newton, che conosceva il calcolo infinitesimale poiché lo aveva lui stesso inventato, al caso di legge qualsiasi di variazione della velocità, e di traiettoria qualsiasi, ed esso è dato, in notazioni moderne, da

$$\mathbf{a}(P,t)=\lim_{\Delta t\to 0}\frac{\mathbf{v}(P,t+\Delta t)-\mathbf{v}(P,t)}{\Delta t}=\frac{d\mathbf{v}(P,t)}{dt} \tag{2'}$$

che, se $x(t)$, $y(t)$, $z(t)$ sono le proiezioni del punto mobile sugli assi del sistema di riferimento, e a_x, a_y, a_z le tre componenti di $\mathbf{a}$ secondo gli assi stessi, dà:

$$a_x=\frac{d^2x}{dt^2} \qquad a_y=\frac{d^2y}{dt^2} \qquad a_z=\frac{d^2z}{dt^2}. \tag{3}$$

Grazie al concetto di accelerazione Galileo poté stabilire la natura dell'effetto sul moto prodotto dall'applicazione di una forza: nella sua opera fondamentale per la meccanica, dal titolo *Discorsi e dimostrazioni matematiche intorno a due nuove scienze attinenti alla Meccanica e i movimenti locali*, egli scrisse: «Qualunque grado di velocità si trovi in un mobile, gli sia per sua natura indelebilmente impresso, purché siano tolte le cause esterne di accelerazione o di ritardamento; il che accade soltanto nel piano orizzontale; infatti nei piani declivi è di già presente una causa di accelerazione, mentre in quelli acclivi [è già presente una causa] di ritardamento: da ciò segue parimenti che il moto sul piano orizzontale è anche eterno; infatti, se è equabile, non scema o diminuisce, né tanto meno cessa» [1638, ed. 1958 p. 249].

L'effetto di una forza non è pertanto la velocità, come ritenevano gli aristotelici, ma l'accelerazione: da osservare, che dalle parole sopra citate risulta chiaramente la *dipendenza* dell'accelerazione dalla forza, e perciò se la forza è costante è costante pure l'accelerazione, ma *non* la proporzionalità di questa a quella.

Si deve appena aggiungere che Galileo, ancora in netta differenziazione dal metodo aristotelico, affermò l'importanza del metodo scientifico sperimentale per lo studio dei fenomeni naturali: già prima di lui, molti, e sopra tutti Bacone, avevano sostenuto la necessità dell'esperienza per detto studio, ma nessuno aveva formulato le regole da seguire per non cadere dal dogmatismo nell'empirismo. Galileo stabilí dette norme, che possono ridursi a queste tre fondamentali:

1) tradurre il fenomeno osservato in numeri che rappresentano i rapporti intercorrenti tra i vari elementi da cui il fenomeno stesso dipende;
2) formulare l'ipotesi di una relazione costante tra questi numeri;
3) ripetere sperimentalmente il fenomeno variandone gli elementi per verificare se la relazione è realmente costante. Se cosí è, questa rappresenta la legge del fenomeno, ossia l'unica verità che lo spirito umano può riconoscervi, poiché la scienza non può conoscere l'essenza del fenomeno, come credeva Bacone, ma solo i rapporti fra i suoi elementi.

Questa filosofia galileiana è bene espressa nel *Proemio ai lettori* dei *Saggi di naturali esperienze* dell'Accademia del Cimento [1667], in cui è scritto: «Or quindi dove non c'è piú lecito metter piede innanzi, non vi à cui meglio rivolgersi che alla fede dell'esperienza, la quale, non altrimenti di chi varie gioie sciolte, e sconnesse cercasse di rimettere ciascuna per ciascuna al suo incastro, cosí ella adattando effetti a cagioni e cagioni ad effetti, se non di primo lancio, come la geometria, tanto fa che PROVANDO E RIPROVANDO le riesce talora di dar nel segno. Conviene però camminar con molto riguardo, che la troppa fede all'esperienza non ci faccia travedere, e ci inganni, ecc.». Galileo, seguendo le regole sopra accennate, osservò il movimento di caduta libera dei gravi, compiendo i suoi esperimenti con corpi diversi, ripetendo e variando le condizioni, e studiò le condizioni da cui tale movimento dipende, deducendo, in base a geniali induzioni fondate sulle sopraddette osservazioni, la legge di caduta libera dei gravi, che in linguaggio moderno si può cosí enunciare: «Un grave abbandonato a se stesso senza velocità iniziale cade lungo la verticale, muovendosi con

accelerazione costante e diretta verticalmente verso il basso, che è la stessa per tutti i corpi».

Per studiare il caso generale dei gravi comunque lanciati, fu guidato da un concetto d'indipendenza degli effetti. Egli intuí che, come nella caduta libera si verifica una variazione di velocità sempre costante, e quindi indipendente dai diversi valori della velocità stessa, cosí debba analogamente avvenire nel caso generale dei gravi lanciati, e l'esperienza confermò questa sua intuizione. Si ha cosí la seconda legge dei gravi: «Un grave lanciato in una direzione qualsiasi e con qualunque velocità iniziale si muove sempre con quella stessa accelerazione costante e diretta verticalmente verso il basso, che si manifesta nei gravi cadenti dalla posizione di quiete». Le sopraddette leggi si traducono, usando il calcolo di Newton e assumendo un sistema di riferimento solidale con la Terra, coll'asse z verticale ed orientato verso il basso, con l'origine nella posizione iniziale del grave, nelle equazioni

$$\ddot{x}=\frac{d^2x}{dt^2}=0 \qquad \ddot{y}=0 \qquad \ddot{z}=g, \tag{4}$$

ove g è una costante, ossia è indipendente non soltanto dal tempo, ma altresí dal grave che cade. L'unica cosa da cui può dipendere g è pertanto il corpo verso cui il moto è diretto, cioè la Terra, e di conseguenza deve essere interpretata come un effetto di una proprietà della Terra, non del corpo che cade: poiché g è una accelerazione, questa proprietà, per il principio galileiano già espresso, deve interpretarsi come una forza esercitata dalla Terra, è cioè la forza di gravità, e g è perciò chiamata accelerazione di gravità. Le (4) risolte dànno

$$\begin{cases} z=\frac{1}{2}gt^2+\dot{z}_0t \\ x=\dot{x}_0t \\ y=0 \end{cases} \tag{5}$$

essendo $\dot{z}_0$ la componente della velocità inizialmente impressa al grave secondo la verticale, $\dot{x}_0$ l'analoga componente secondo la retta scelta come asse x nel piano orizzontale, mentre si è preso $\dot{y}_0=0$. L'equazione della traiettoria del grave si ottiene dalle (5) eliminando t, e risulta definita dalla

$$z=\frac{g}{2\dot{x}_0^2}x^2+\frac{\dot{z}_0}{\dot{x}_0}x \tag{6}$$

e pertanto il grave lanciato da un punto qualsiasi con velocità $\dot{x}_0\mathbf{i}_x+\dot{z}_0\mathbf{i}_z$ pure qualsiasi descrive come traiettoria la parabola del piano (x, z) data dalla (6).

I risultati ora dati, al pari delle leggi sopra ricordate, valgono solo se il moto avviene nel vuoto, come già Galileo fece osservare; inoltre essi sono stati ricavati prendendo un sistema di riferimento solidale con la Terra, che non è un sistema di riferimento assoluto, anche secondo Galileo, e perciò sono validi nei limiti nei quali è lecito trascurare l'influenza della rotazione terrestre sul fenomeno in studio, come sarà precisato in seguito. È opportuno peraltro indicare

già qui la relazione che lega le accelerazioni di uno stesso punto mobile P rispetto a due riferimenti diversi S ed S', noto essendo il moto ad esempio di S' rispetto a S (teorema di composizione delle accelerazioni), sempre nell'ipotesi di esistenza del concetto di tempo assoluto. Questa relazione si ricava da quella già data relativa al principio di composizione delle velocità, e precisamente: siano $\mathbf{a}_a$ e $\mathbf{a}_r$ rispettivamente le accelerazioni di P nei moti di P rispetto a S e di P rispetto a S'; si consideri il moto di S' rispetto a S composto di un moto traslatorio di velocità uguale alla velocità dO'/dt dell'origine O' di S' (e pertanto producente in tutti i punti solidali con S' la medesima velocità dO'/dt), e di un moto rotatorio $\boldsymbol{\omega}$ attorno ad un asse passante per O' (e pertanto producente in ogni punto P' di S' la velocità $\boldsymbol{\omega}\wedge(P'-O')$). È allora

$$\mathbf{a}_a=\mathbf{a}_r+\mathbf{a}_t+\mathbf{a}_c \tag{7}$$

in cui

$$\mathbf{a}_r=\frac{d^2x'}{dt^2}\mathbf{i}_{x'}+\frac{d^2y'}{dt^2}\mathbf{i}_{y'}+\frac{d^2z'}{dt^2}\mathbf{i}_{z'},$$

è l'accelerazione di P nel moto relativo a S'; $\mathbf{a}_t=(d^2O'/dt^2)+(d\boldsymbol{\omega}/dt)\wedge(P-O')+$ $+\boldsymbol{\omega}\wedge\{\boldsymbol{\omega}\wedge(P-O')\}$ è l'accelerazione del punto P (rispetto a S) pensato solidale con S' (accelerazione di trascinamento); e infine $\mathbf{a}_c=2\boldsymbol{\omega}\wedge\mathbf{v}_r$ è l'accelerazione di Coriolis (o complementare): la (7), nella forma ora indicata, apparve per la prima volta in una memoria del «Journal de l'Ecole Polytechnique» nel 1836 ad opera di Gustave-Gaspard Coriolis (e da ciò il nome dato ad $\mathbf{a}_c$).

1.2. Da Newton ad Einstein.

Se Galileo iniziò, Newton continuò e completò la creazione della dinamica come scienza del moto: la sua opera fondamentale in argomento è *Philosophiae Naturalis Principia Mathematica*, stampata il 5 luglio 1686 e tradotta in inglese da Andrew Motte nel 1729. Egli introdusse in modo esplicito un altro concetto fondamentale, quello di massa, come concetto che è distinto da quello di peso, proporzionale a questo, ma che, a differenza di questo, ha carattere assolutamente intrinseco, dipendente dalla natura materiale del corpo, ma scevro da qualsiasi influenza esterna locale: per questo suo carattere assolutamente intrinseco la massa di un sistema qualsiasi si conserva e *non* dipende dal suo stato di quiete o di moto. Definí quindi: la quantità di moto, come il prodotto della massa per la velocità (ed è singolare che nelle lingue latine questa grandezza abbia conservato la denominazione sopraddetta data da Newton, mentre nella lingua inglese l'espressione corrispondente *quantity of motion* usata dal Motte sia poi stata sostituita da *momentum*); la forza d'inerzia (*vis insita*, o *vis inertiae*) come la forza con la quale la massa del corpo si oppone, resiste, ad ogni azione che tende a variare il suo stato di quiete o di moto; la forza applicata (o impressa) come l'«azione esercitata sul corpo al fine di mutare il suo stato di quiete o di moto rettilineo uniforme. Questa forza consiste nell'azione in quanto tale, e, cessata l'azione, non permane nel corpo. Infatti un corpo persevera in ciascun nuo-

vo stato per la sola forza di inerzia. La forza impressa ha varie origini: l'urto, la pressione, e la forza centripeta» [1713, trad. it. p. 95].

Newton enuncia quindi i tre assiomi, o le tre leggi del moto, che costituiscono il fondamento di tutta la meccanica newtoniana:

PRIMA LEGGE (o legge d'inerzia). *Ogni corpo persevera nel suo stato di riposo, oppure di moto rettilineo uniforme, a meno che non sia costretto a cambiare tale stato da forze agenti su di esso* (nella versione originale in latino: «Corpus omne perseverare in statu suo quiesciendi vel movendi uniformiter in directum, nisi quatenus a viribus impressis cogitur statum illum mutare»).

SECONDA LEGGE (o legge di azione). *Il cambiamento della quantità di moto è proporzionale alla forza motrice impressa ed avviene nella stessa direzione della retta secondo la quale agisce la forza impressa* (nella versione latina: «Mutationem motis proportionalem esse vi motrici impressae, et fieri secundum lineam rectam qua vis illa imprimitur»).

TERZA LEGGE (o legge di azione e reazione). *Ad ogni azione corrisponde sempre una reazione uguale e contraria, ossia, le forze reciproche che si scambiano due corpi sono sempre uguali ed opposte* (nella versione latina: «Actioni contrariam semper et aequalem esse reactionem: sive corporum duorum actiones in se mutuo semper esse aequales et in partes contrarias dirigi»).

A base dei tre assiomi sopra specificati stanno i postulati di tempo assoluto e di spazio (e sistema di riferimento) assoluto, che non traggono la loro origine da una evidenza sperimentale, cosí che lo stesso Newton, nello *scholium* al termine del capitolo primo, relativo alle *Definizioni* dice: «Le parti dello spazio immobile, in cui i corpi veramente si muovono, non cadono sotto i sensi» [*ibid.*, p. 110] e ancora: «Per quanto sia possibile che un qualche corpo, nelle regioni delle stelle fisse, o anche piú lontano, sia in quiete assoluta, tuttavia, dalla posizione fra loro dei corpi nelle nostre regioni, non si può sapere se qualcuno di questi conserva o no una data posizione rispetto a quel corpo tanto lontano» [*ibid.*, p. 107]. La critica di questi postulati, e la loro modificazione, dalla quale deriveranno pure variazioni nei concetti di massa, per quanto riguarda il suo carattere assolutamente intrinseco, e di forza, almeno per quanto si riferisce alla sua interpretazione nei fenomeni celesti, sarà sviluppata due secoli piú tardi da Einstein e di essa sarà fatto cenno in seguito: per ora, quello che importa dire è che per il periodo sopra accennato non solo tutti i risultati dedotti con la meccanica newtoniana apparirono in ottimo accordo con i risultati sperimentali, ma essa ha portato pure a mirabili progressi in fisica, in astronomia, meccanica applicata. Del resto le variazioni cui sopra si è fatto cenno, ed altre ancora, che pure saranno in seguito accennate, hanno sí indicato limitazioni al campo di validità della meccanica newtoniana, ma hanno anche mostrato come questa sia pienamente applicabile a un grandissimo numero di problemi, che interessano la scienza e la tecnica.

La seconda legge di Newton si traduce nella celebre relazione

$$\mathbf{F} = m\mathbf{a} \tag{8}$$

in cui **F** è il risultante delle forze applicate al punto materiale mobile, la cui massa è m e la cui accelerazione è **a**. La (8) è valida solo se il moto è riferito a un sistema di riferimento assoluto, e pertanto **a** è l'accelerazione assoluta; ma dal teorema di composizione delle accelerazioni (5) risulta subito che se le leggi meccaniche sono valevoli per un dato sistema di riferimento, esse lo sono pure per qualsiasi altro sistema in traslazione uniforme rettilinea rispetto al primo: è questo il principio oggi universalmente chiamato, secondo la proposta di Einstein nella sua prima opera sulla relatività, principio galileiano di relatività, mentre sono indicate come terne inerziali galileiane le terne assolute, rispetto alle quali la (8) è valida.

Newton indubbiamente applicò il metodo galileiano di collegare lo studio teorico a quello sperimentale; peraltro la presentazione dei risultati da lui ottenuti in forma assiomatica, come fa nei suoi *Principia*, non agevola la comprensione del modo col quale essi sono stati raggiunti. In effetti, egli, benché avesse, prima ancora di creare la scienza del moto, inventato il calcolo infinitesimale (la teoria delle flussioni esposta in *Methodus fluxionum et seriarum infinitarum, cum eiusdem applicationes ad curvarum geometriam*), nei *Principia* fa uso soltanto di costruzioni geometriche, secondo lo stile classico dei Greci. Max Born, nella sua *Natural Philosophy of Cause and Chance* [1949], dice a questo proposito: «Non so se egli avesse realmente scoperto quei risultati con l'aiuto del calcolo infinitesimale; mi sembra incredibile il contrario. Egli era indubbiamente desideroso di evitare nuovi metodi matematici allo scopo di conformarsi al gusto dei suoi contemporanei. Ma è anche noto che egli amava celare sotto un rivestimento le proprie idee reali. Questa stessa tendenza si riscontra in Gauss e in altri grandi matematici, ed è sopravvissuta fino ai nostri giorni, con gran danno per la scienza» (trad. it. p. 166). (La *Natural Philosophy* è del 1949; che cosa direbbe Max Born oggi?)

Comunque, Newton riconosce la priorità di Galileo per le prime due leggi, almeno per quanto riguarda la loro applicazione al moto dei gravi, in quanto dice: «Per mezzo delle prime due leggi e dei primi due corollari Galileo trovò che la caduta dei gravi è proporzionale al quadrato del tempo, e che il moto dei proiettili avviene secondo una parabola» [1713, trad. it. p. 126]. Dalle prime due leggi di Keplero egli ricavò quindi che l'accelerazione del centro del pianeta P è sempre diretta verso il centro del Sole S e varia in ragione inversa del quadrato della distanza r_P da questo:

$$\mathbf{a}_P = \frac{\mu}{r_P^2}\boldsymbol{\rho}_{PS} \tag{9}$$

se $\boldsymbol{\rho}_{PS}$ è il versore della retta PS orientata da P verso S, e μ una costante, che può differire da pianeta a pianeta. Dalla terza legge peraltro si ricava, che, se A è il semiasse maggiore dell'orbita ellittica e T il periodo di rivoluzione, è

$$\frac{A^3}{T^2} = \text{cost} = \frac{\mu}{4\pi^2}. \tag{9'}$$

L'estensione del procedimento sopra accennato, prima al moto dei satelliti attorno ai rispettivi pianeti, per i quali pure valgono con buona approssimazione le leggi di Keplero (come è già stato detto per il caso Luna-Terra) e, poi, passo questo certo molto ardito e sorprendente che provocò la piú grande ammirazione nei contemporanei di Newton e nelle successive generazioni, la inclusione dei corpi terrestri nelle leggi ricavate per i corpi celesti (il che comporta l'ammettere che la forza esterna agente sul grave in caduta o sulla Luna in moto lungo la sua orbita abbiano la stessa origine: l'attrazione della Terra sui corpi materiali) portarono alla

LEGGE DELLA GRAVITAZIONE UNIVERSALE (Newton). *Nell'Universo due punti materiali qualsiasi si attraggono reciprocamente con una forza diretta secondo la loro congiungente, proporzionale direttamente alle loro masse m e m', e inversamente al quadrato della loro distanza r: $K\, mm'/r^2$.*

La costante K è chiamata costante di attrazione (o costante del Gauss) e misura la mutua attrazione fra due masse unitarie all'unità di distanza: essa fu determinata per la prima volta da Henry Cavendish nel 1797, e secondo le piú recenti esperienze ha il valore in unità CGS $6{,}7 \times 10^{-8}$.

È da osservare che, ammessa la legge di Newton della gravitazione, le leggi di Keplero non possono piú considerarsi come rigorosamente valide, perché queste si verificano solo quando i corpi che si attraggono sono due soltanto e il corpo centrale è fisso (ossia può prendersi come origine di un sistema inerziale); si può però riconoscere che sia le leggi di Keplero, sia gli altri risultati sperimentali acquisiti sono compatibili con la legge di gravitazione entro accettabili limiti di approssimazione, o, meglio ancora, che ogni qualvolta si sono osservati divari tra i risultati sperimentali e i risultati calcolati nell'ipotesi che i corpi interagenti fossero solo due, e il sistema di riferimento connesso con uno dei corpi stessi fosse un sistema inerziale, si è constatato che detti divari erano prodotti da azioni perturbatrici conseguenti all'una o all'altra, o ad entrambe le sopraddette ipotesi, ed essi venivano eliminati coll'applicazione piú corretta delle leggi newtoniane: la scoperta del pianeta Nettuno, avvenuta nel 1846, e quella di Plutone, fatta nel 1930, ed entrambe avvenute per questa via, sono tra le piú brillanti testimonianze della validità della legge di Newton della gravitazione universale.

Si può ricavare facilmente l'ordine di grandezza dell'errore corrispondente ad esempio all'ammettere che sia inerziale la terna solidale con la Terra nella determinazione della legge di caduta dei gravi, se si considera che dal teorema di Coriolis (9), detta $\mathbf{a}_r$ l'accelerazione del grave nel moto relativo alla Terra, ed essendo $\boldsymbol{\omega}$ la rotazione del sistema di riferimento a questa solidale, l'equazione del moto relativo è data, invece che dalle (4), dalla

$$\mathbf{a}_r = \mathbf{g} - 2\boldsymbol{\omega} \wedge \mathbf{v}_r. \tag{10}$$

Preso un sistema di assi, di cui l'origine O coincide con la posizione iniziale del grave, l'asse z sulla linea d'azione della forza peso in O (verticale del luogo) orientata verso il basso, l'asse x nel piano meridiano terrestre per O, orientato

verso nord, e pertanto l'asse y perpendicolare a detto piano meridiano e orientato verso est, dalla (10) si ricava, per velocità iniziale nulla,

$$x = 0 \quad y = \frac{1}{3}\omega g(\cos\gamma)t^3 \qquad z = \frac{1}{2}gt^2 \tag{11}$$

essendo γ l'angolo acuto formato da $\mathbf{g}$ col piano equatoriale (latitudine geoidica). Si riconosce pertanto, che essendo $\cos\gamma > 0$ è $y > 0$, e pertanto un grave liberamente cadente senza velocità iniziale *non* segue la verticale del luogo, ma devia dal filo a piombo verso est: per $t = 1$ s, essendo $\omega = 2\pi/86164\ \mathrm{s}^{-1}$, alla latitudine di Roma risulta $y = 0{,}18$ mm. Questa deviazione orientale dei gravi fu prevista teoricamente da Newton ed è in accordo col risultato sperimentale di Tadini (1795).

Un'altra osservazione deve essere fin d'ora fatta: le assunzioni a base della deduzione della legge gravitazionale di Newton implicano che la massa inerte m_i, ossia quella che misura il rapporto tra la forza impressa a un corpo materiale e la corrispondente accelerazione prodotta, sia uguale alla massa pesante m_p dello stesso corpo, ossia al coefficiente cui è proporzionale (secondo la costante universale K sopra definita) la forza di attrazione esercitata da detto corpo sull'unità di massa di un altro corpo qualsiasi, a distanza unitaria dal primo: ne consegue che il rapporto delle masse di due corpi, ricavato come rapporto delle accelerazioni ad essi impresse da una medesima forza, è uguale al rapporto dei pesi ricavato con una bilancia. È chiaro che la differenza tra le due determinazioni consiste in questo, che la prima è del tutto indipendente dalla forza di gravità, mentre la seconda è dovuta all'esistenza di tale forza. Ora, che le due determinazioni diano il medesimo risultato non è a priori evidente: è un risultato dell'esperimento, e non del ragionamento. A questo riguardo Max Born [1949] scrive: « Né lo stesso Newton, né le molte generazioni di fisici e di astronomi venute dopo di lui hanno fatto molta attenzione alla legge espressa dalla

$$m_i = m_p. \tag{12}$$

« Le osservazioni astronomiche non lasciavano dubbi che essa fosse corretta, e le osservazioni terrestri (con pendoli adatti) ne dimostravano la validità con estrema accuratezza (esperimenti di Eötvos e altri). Dovevano trascorrere due secoli prima che Einstein scorgesse il problema fondamentale contenuto nella semplice equazione (12), e costruisse su di essa il colossale edificio della sua teoria sulla relatività generale » (trad. it. p. 27).

È ancora interessante osservare che nello *scholium* al capitolo primo dei *Principia* si ritrova la legge aristotelica contenente in nuce il principio dei lavori virtuali, già prima citata. In esso infatti è detto: « Identica è la norma di tutte le macchine. L'efficacia e l'uso di esse consiste in questo solo, che diminuendo la velocità aumentiamo la forza e viceversa: per cui, in ogni genere di strumenti adatti viene risolto il problema, *muovere un dato peso con una data forza*, e superare con una data forza un'altra data resistenza » [Newton 1713, trad. it. p. 135]. Piú avanti egli poi enuncia una proposizione che è in effetti il teorema dell'energia (in meccanica): « Se noi stimiamo l'azione dell'agente dal prodotto della sua forza per la velocità, e analogamente la reazione di tutto ciò che si oppone al moto dal pro-

dotto delle velocità delle sue singole parti per le forze di resistenza ad esse applicate dovute ad attrito, coesione, peso e accelerazione, l'azione e la reazione saranno sempre uguali l'una all'altra». Newton non dà nessun nome a questo teorema, mentre la denominazione di «forza viva» assegnata alla grandezza $mv^2/2$, per una generica particella di massa m e velocità v, è dovuta a Leibniz (in realtà Leibniz introdusse come forza viva la mv^2).

La dinamica newtoniana, nonostante i grandi successi ottenuti nei problemi ai quali fu applicata dal suo autore, fu per diversi anni avversata da parecchi contemporanei di Newton: anzitutto, la forza di attrazione newtoniana appariva *non* come una forza che *spingesse* i pianeti, ma come una forza che agiva a distanza dal corpo da cui emanava su quello cui era applicata senza alcun intermediario tra di essi, ed agiva non nel senso del moto, ma in modo da deviare questo moto; la natura della forza stessa appariva poi del tutto oscura. Per questo, ancora venticinque anni dopo la pubblicazione dei *Principia*, nella stessa Inghilterra alla cosmologia di Newton era preferita quella di Descartes, nella quale i pianeti erano sospinti sulle loro orbite da vortici di un fluido celeste, allo stesso modo come si vedono pezzi di legno fatti ruotare dai vortici di un corso d'acqua, e nel resto d'Europa la teoria vorticosa di Descartes ebbe vita ancora piú lunga, cosí che Voltaire, che visitò Londra proprio nel 1727 (anno della morte di Newton), scrisse che quanto Newton morí il numero dei suoi seguaci non era superiore a venti. Si può dire che solo nella seconda metà del secolo XVIII il sistema newtoniano fu universalmente accettato.

Newton era ben conscio delle difficoltà insite nella sua teoria: in una lettera a Richard Bentley del 1692 egli si oppone fortemente all'idea che la gravitazione fosse una proprietà intrinseca o innata della materia, cosí come contrasta vivamente l'idea di un'azione a distanza. Egli scrive: «Voi qualche volta parlate della gravità come di una proprietà essenziale e inerente alla materia. Prego, non attribuite a me queste idee, perché io non pretendo di conoscere la causa della gravità». In un'altra lettera del 1693 egli dice: «È inconcepibile che materia bruta, inanimata possa, senza la mediazione di qualcosa altro, che non sia materiale, agire e operare sopra altra materia senza mutuo contatto»; e ancora: «Che la gravità sia innata, insita ed essenziale nella materia, di guisa che un corpo possa agire sopra un altro a distanza attraverso il vuoto, senza la mediazione di qualche cosa d'altro, dal quale e attraverso il quale la loro mutua azione e forza possa essere convogliata dall'uno allo altro corpo è certo una cosí grande assurdità che io credo, che nessun uomo, che abbia una facoltà di pensare in argomenti filosofici, lo potrà mai ammettere. La gravità deve essere causata da un agente, che agisce costantemente secondo certe leggi, ma che questo agente sia materiale o immateriale, lo lascio alla considerazione dei miei lettori».

Per quanto poi riguarda la natura della forza d'attrazione tra i corpi celesti e terrestri nello *scholium* generale al termine dei *Principia mathematica* [1713] egli scrive: «Fin qui ho spiegato i fenomeni del cielo e del nostro mare mediante la forza di gravità, ma non ho mai fissato la causa della gravità... In verità non sono ancora riuscito a dedurre dai fenomeni la ragione di queste proprietà della gravità, e non invento ipotesi» (trad. it. p. 795).

Il problema della natura della forza di gravità fu di nuovo affrontato, a piú di due secoli di distanza da Newton, da Einstein nella sua teoria della relatività generale, della quale sarà fatto cenno piú avanti e per la quale l'azione gravitazionale appare come una conseguenza della modificazione strutturale del continuo quadridimensionale spazio-tempo, prodotta dalla materia stessa. Eddington, nella *Nature of the Physical World*, dice: «La legge di gravitazione di Einstein è regolata da una grandezza geometrica, la curvatura, in contrasto con la legge di Newton, che è regolata da una grandezza meccanica, la forza» [1928, ed. 1929 p. 133]; e certamente la riduzione della gravità a un effetto geometrico della curvatura dello spazio-tempo è un concetto rivoluzionario di grande portata per le conseguenze cui porta, per quanto, peraltro, fu osservato, esso corrisponda piú a un modello matematico del fenomeno che non a una spiegazione fisica di questo. Ricordando le diffidenze, che accompagnarono l'opera di Einstein, non c'è certo da stupirsi se uguale sorte toccò sul principio anche all'opera di Newton; comunque, dopo un primo periodo di diffidenza, le teorie di Newton ebbero un successo travolgente, e per alcuni decenni esse regnarono incontrastate, in quanto le numerosissime ed importantissime ricerche, che furono fatte in detta epoca, riguardano non l'analisi critica dei concetti e delle leggi fondamentali su cui si basano, ma essenzialmente l'indagine matematica delle equazioni del moto, e per questa si citano qui le opere fondamentali di D'Alembert, di Maupertuis, di Lagrange, di Eulero, di Laplace, di Gauss, di Hamilton, di Hertz e di molti altri.

Non è questo il luogo per indicare neppure sommariamente la mirabile sistemazione che da queste opere fu data alla meccanica newtoniana, e per quanto riguarda la deduzione di principî generali, che permettono di esprimere le equazioni del moto dei corpi in una forma particolarmente suggestiva e significativa, e per quanto si riferisce ai metodi matematici per la risoluzione delle equazioni che da detti principî possono essere ricavate. Si ricorda qui soltanto per la sua sinteticità e la sua potenza il principio dell'azione stazionaria, o principio di Hamilton [1834]: esso si esprime, per un sistema dinamico discreto soggetto a forze conservative (che derivano cioè da un potenziale), per il quale la termodinamica può essere ignorata, con la relazione

$$\delta\left[\int_{t_1}^{t_2} L\,dt\right]=0 \tag{13}$$

essendo L la cosiddetta funzione lagrangiana definita dalla

$$L=T-V \tag{14}$$

se T è l'energia cinetica (o forza viva) del sistema e V la sua energia potenziale; la grandezza fra parentesi nella (13) è chiamata azione, e il significato della (13) è:

Si considera il moto (chiamato «moto variato sincrono»), nel quale i punti del sistema hanno, in ogni medesimo istante t, coordinate $x_r+\delta x_r$ infinitamente vicine alle coordinate x_r che corrispondono al moto effettivo (e compatibili coi vincoli ad essi eventualmente posti), e si calcola la variazione che l'azione presenta tra detto moto variato sincrono e il moto effettivo nell'ipotesi che le posizioni del sistema relative agli istanti t_1 e t_2 siano le stesse per entrambi i moti:

la (13) esprime che detta variazione è nulla, qualunque sia il moto variato considerato, e qualunque siano gli istanti t_1 e t_2. Il moto effettivo, per il principio di Hamilton, viene cosí ad essere caratterizzato, fra tutti i moti possibili del sistema fra le medesime posizioni estreme, come quello che rende stazionaria l'azione, e da ciò il nome sopra ricordato per il principio stesso. Questo poi può essere generalizzato al caso di un sistema continuo, per il quale non si abbia una produzione irreversibile di entropia, con la

$$(13') \qquad \delta\left[\int_{t_1}^{t_2}(T-V-U)\,dt\right]=0$$

in cui U è l'energia interna del sistema stesso.

Le opere cui sopra si è fatto cenno portarono la meccanica newtoniana a un tal grado di perfezione logica, e a una tale potenza nel descrivere e nel predire esattamente le osservazioni sperimentali, da indurre nella convinzione che ogni scienza dovesse essere ridotta alla meccanica, o almeno modellarsi sulla meccanica di Newton. Peraltro, la continua estensione delle conoscenze scientifiche, e la molteplicità dei problemi cui questa ha portato, dopo vani tentativi d'interpretare tutti i fenomeni naturali partendo da criteri meccanicistici che condussero ad inventare sostanze artificiose come il calorico, i fluidi elettrico e magnetico, l'etere, determinarono a partire dalla seconda metà del secolo scorso una decadenza di tale interpretazione meccanicistica, e portarono a una profonda revisione degli stessi concetti fondamentali della meccanica.

Questa revisione, che in effetti condusse a una vera e propria rivoluzione nel pensiero scientifico, è dovuta ad Albert Einstein [1905*a*], che pubblicò negli «Annalen der Physik» una nota basata sui seguenti postulati:

1) la velocità della luce è la stessa in tutti i sistemi di riferimento in moto traslatorio rettilineo uniforme gli uni relativamente agli altri;
2) tutte le leggi della natura sono le stesse in tutti i sistemi di riferimento in moto relativo traslatorio rettilineo uniforme.

Quali siano le conseguenze che da detti postulati derivano e come esse siano dedotte da questi è indicato in altro articolo di questa stessa *Enciclopedia*, e perciò qui ci si deve limitare a un breve cenno solo a quei risultati che piú strettamente si ricollegano a quanto prima è stato detto.

La prima delle proposizioni sopra enunciate esprime una proprietà dimostrata sperimentalmente dapprima da Michelson e Morley, i quali con un interferometro di Michelson orientato in vari modi rispetto al moto della Terra avevano trovato che se una sorgente di luce si muove rispetto all'etere, le cui vibrazioni elettriche si identificavano con quelle della luce, e se un osservatore si muove rispetto al medesimo, la velocità di propagazione della luce nel vuoto è sempre $c=2{,}9979245\times 10^{10}$ cm/s. In epoca recente questo risultato è stato confermato con accuratezza molto maggiore che nel passato da Kenneth Brecher [1977] del Massachusetts Institute of Technology, il quale fece uso delle sorgenti a raggi X regolarmente pulsanti in sistemi binari di stelle.

Il secondo dei postulati sopra indicati, che è noto come principio di relatività ristretta, è una generalizzazione del principio galileiano di relatività, in quanto esso si estende ad ogni legge fisica: esso equivale ad affermare che non è possibile distinguere un sistema di riferimento rispetto ad un altro muoventesi di moto traslatorio rettilineo con velocità costante rispetto al primo, perché ogni esperienza eseguita nel primo e nel secondo sistema dà lo stesso risultato.

I risultati, che da questi postulati derivano, mettono in evidenza come nella nuova meccanica, su di essi basata, un orologio muti di ritmo, e un regolo muti di lunghezza, secondo leggi di variazione rigorosamente determinate, nel passaggio dallo stato di quiete a quello di moto, e precisamente, se dt è la durata elementare di un fenomeno, misurato da un osservatore $(O;\,S)$, e dl la distanza elementare fra due punti dello spazio geometrico tridimensionale riferito al sistema di riferimento S, ciò che rimane invariato nel passaggio da S a un altro riferimento S' in moto traslatorio rettilineo uniforme rispetto al primo non è la durata dt né la distanza spaziale dl, bensí la grandezza

$$(15) \qquad ds=\sqrt{c^2(dt)^2-(dl)^2}=\sqrt{c^2(dt)^2-dx_1^2-dx_2^2-dx_3^2}$$

che misura, in una metrica opportuna, la distanza tra due punti (chiamati «eventi») dello spazio quadridimensionale (cronotopo). Questo porta ad abbandonare le concezioni di spazio assoluto e di tempo assoluto, ma è bene osservare subito che dalle stesse formule che dànno il risultato sopra indicato appare che detti concetti conservano piena validità se i valori delle velocità dei punti materiali in moto sono trascurabili rispetto alla velocità della luce.

Un altro concetto fondamentale connesso col moto, che nella meccanica relativistica subisce una revisione profonda, è quello corrispondente alla massa, che perde il carattere assolutamente intrinseco che esso ha nella meccanica newtoniana ed appare invece dipendente dal moto. In effetti, esprimendo il principio di conservazione della massa e della quantità di moto rispetto ai due osservatori solidali ai riferimenti S e S' prima definiti, si ottiene che se m_0 è la massa di una particella P quale è misurata da un osservatore rispetto al quale essa è in quiete (chiamata massa a riposo, o massa propria), la sua massa misurata dall'osservatore rispetto al quale ha una velocità $\mathbf{u}(u_1,\,u_2,\,u_3)$ è

$$(16) \qquad m=\frac{m_0}{\sqrt{1-\dfrac{u^2}{c^2}}}.$$

Questo riconoscimento del carattere non intrinseco della massa porta subito un problema: la prima legge di Newton è senz'altro ammessa nella meccanica einsteiniana, ma per quanto si riferisce alla seconda legge è da osservare che mentre per la meccanica newtoniana è indifferente scrivere $m(d\mathbf{u}/dt)=\mathbf{F}$, oppure $(d/dt)(m\mathbf{u})=\mathbf{F}$, lo stesso non si ha in quella di Einstein: in questa si prende

$$(17) \qquad \frac{d}{dt}\frac{m_0\mathbf{u}}{\sqrt{1-\dfrac{u^2}{c^2}}}=\mathbf{F}$$

e questa relazione, in quanto permette di confrontare i diversi effetti prodotti da una medesima causa del moto applicata a diversi corpi, è assai piú che una pura definizione di forza. L'assunzione della (17) non è però conseguenza di un puro ragionamento, e la sua validità è da ricercarsi essenzialmente nella coincidenza dei risultati che da essa si deducono coi risultati degli esperimenti. Da osservare poi che imponendo la condizione di invarianza relativistica alla legge rappresentata dalla (17) rispetto ai riferimenti S ed S', si deduce che mentre nella meccanica newtoniana la misura della forza agente sopra una data particella è la stessa per tutti i sistemi di riferimento in moto relativo traslatorio rettilineo uniforme, lo stesso non accade nella meccanica einsteiniana: però, ancora, il divario tende ad annullarsi per $c \to \infty$. Si riconosce inoltre che se in uno stesso sistema di riferimento due particelle hanno diverse velocità e sono soggette ad uguali forze, in ogni altro sistema di riferimento le stesse particelle appaiono soggette a forze diverse. Un risultato analogo vale nel caso di due particelle soggette in S a due forze uguali e contrarie, e perciò appare che la terza legge di Newton non ammette un'immediata estensione alla meccanica relativistica. Infine: moltiplicando scalarmente i membri della (17) per **u** si ricava l'equazione che esprime il teorema dell'energia nella meccanica relativistica

$$\frac{d}{dt}(mc^2) = \frac{d\mathfrak{L}}{dt} \tag{18}$$

in cui $d\mathfrak{L}/dt$ è il lavoro svolto nell'unità di tempo dalla forza **F**, ossia la sua potenza. Ora la (18) è suscettibile di un'interpretazione che conferisce all'equazione stessa, dal punto di vista concettuale, un valore enorme: da essa appare infatti che una particella di massa m ha l'energia

$$E = mc^2 \tag{19}$$

e pertanto la *massa rappresenta energia*, e viceversa l'*energia possiede massa*. I due concetti fondamentali della meccanica di Newton, massa ed energia, vengono a ridursi a un unico concetto fondamentale, quello dell'energia. È illuminante a questo riguardo quanto dice Einstein in proposito: «Come mai il fatto che l'energia possiede massa e che la massa rappresenta energia è rimasto cosí a lungo allo scuro?... La causa... è il piccolissimo rapporto di scambio fra materia ed energia... La quantità di calore necessaria per convertire trentamila tonnellate di acqua in vapore non peserebbe piú di un grammo circa!» [Einstein e Infeld 1938, trad. it. p. 207].

I pochi cenni, ai quali qui ci si è limitati, sono già sufficienti perché si possa capire la ragione per la quale la meccanica einsteiniana al suo apparire, e per qualche decennio, incontrò molte diffidenze e molte opposizioni; essa finí però per imporsi, perché tutte le proprietà da essa dedotte sono state mirabilmente confermate dagli esperimenti. Questo è avvenuto per la proprietà della dilatazione dei tempi nei corpi in moto, che da qualcuno è presentata come il paradosso dei gemelli, ed ha trovato conferma ad esempio nella misura del decadimento dei muoni o mesoni μ, scoperti nella radiazione cosmica; per l'equivalenza tra massa ed energia, dimostrata ad esempio dall'energia che si libera nel

processo di fusione di due nuclei di deuterio ($_1H^2$) portanti alla generazione delle particelle α (nucleo dell'isotopo dell'elio $_2He^4$), che è alla base delle ricerche per creare centrali nucleari all'idrogeno; per la variabilità della massa con la velocità, bene constatata nella costruzione dei grandi acceleratori di particelle.

La teoria della relatività ristretta (o speciale), se porta ad una profonda revisione dei concetti fondamentali di spazio, di tempo e di massa, conserva ancora parecchi concetti fondamentali della meccanica newtoniana, e precisamente quelli di moto assoluto e di riferimenti inerziali, rispetto ai quali soltanto la seconda legge di Newton, e quindi anche la (17), è valida; afferma che tutte le leggi della natura, e non soltanto quelle meccaniche, sono invarianti e valevoli per tutti i sistemi inerziali, ma lascia senza risposta la questione fondamentale: esiste un sistema inerziale? Per quanto riguarda la teoria della gravitazione, lascia invariati i punti interrogativi sulla natura dell'azione a distanza e sulla propagazione istantanea dell'azione gravitazionale.

Le difficoltà insormontabili (o almeno ritenute tali nella prima metà del presente secolo), attinenti all'esistenza di un sistema inerziale e di un moto assoluto, di guisa che si ha il paradosso che si conoscono le leggi della meccanica ma non si sa rispetto a quale riferimento esse siano valide, portarono Einstein a concepire una nuova teoria (teoria della relatività generale), nella quale l'invarianza delle leggi fisiche ha luogo non soltanto per ogni trasformazione di Lorentz, corrispondente al passaggio da un osservatore a un altro in moto traslatorio rettilineo uniforme rispetto al primo, ma anche rispetto ad un *qualunque* cambiamento di riferimento spazio-temporale, e perciò anche rispetto a quelle trasformazioni che definiscono il passaggio da un osservatore a un altro comunque mobile rispetto al primo.

Si soddisfa nel modo piú naturale detta condizione d'invarianza esprimendo le leggi fisiche con relazioni tensoriali nello spazio-tempo, di guisa che le equazioni della relatività generale hanno l'aspetto di equazioni strutturali che ricordano quelle che definiscono lo stato di deformazione e di tensione in una struttura continua elastica. Si rimanda all'apposito articolo di questa stessa *Enciclopedia* per una conoscenza delle complesse trattazioni che sono state sviluppate prima da Einstein, e quindi, oltre che da lui, da Eddington, Kaluza e Klein, Weyl e molti altri, per attuare rigorosamente il programma della relatività generale. Importa però osservare qui che se uno dei principali motivi ispiratori di detta teoria è stato quello di fare scomparire quelli che Einstein chiamò «i due ingombranti fantasmi dello spazio assoluto e del sistema inerziale» [*ibid.*, p. 248], l'importanza, l'interesse e la validità di essa trascendono di gran lunga questo scopo, cosí che se anche si riuscisse in qualche modo ad individuare il sistema di riferimento al quale si potesse attribuire un carattere di speciale privilegio, non per questo tali importanza, interesse e validità sarebbero diminuiti. Questo va detto, perché, in effetti, un gruppo di sperimentatori di Berkeley (1977), usando un aereo U-2, sembra aver rilevato recentemente il moto della Terra rispetto al sistema delle microonde cosmiche, che si crede costituiscano quello che rimane della fase primordiale della genesi dell'Universo. Importa poi pure osservare che la meccanica newtoniana è ancora contenuta come caso li-

mite in quella einsteiniana, ma questa, anche nelle condizioni limiti, contiene molto di piú, perché, applicata ad esempio alla cosmologia, da essa si deriva non soltanto l'equazione del moto, o se si vuole la legge d'inerzia, ma anche l'equazione cui deve soddisfare la funzione U, che nella meccanica di Newton ha l'ufficio di potenziale gravitazionale. Quella poi che nella meccanica di Newton è interpretata come forza di gravità appare essere una conseguenza della deformazione del cronotopo, prodotta dalla materia stessa: la presenza di energia, di qualsiasi natura, deforma ed incurva il cronotopo, e la sua influenza sul moto si manifesta attraverso tale deformazione; e ancora, quell'equivalenza tra massa pesante e massa inerte, che nella vecchia teoria appariva puramente casuale ed affermata soltanto dall'esperimento, appare nella nuova teoria come un principio insito nella teoria stessa.

Nella condizione limite, come si è detto, non vi è alcuna differenza, a parte la diversa interpretazione cui si è accennato, tra la legge del moto secondo la teoria newtoniana e quella einsteiniana: un pianeta ad esempio descrive sempre un'ellisse di cui il Sole è un fuoco. Ma se ci si vale delle soluzioni rigorose delle equazioni gravitazionali si ricavano nuove deduzioni, che non possono essere tratte dalla legge newtoniana; tali ad esempio sono: lo spostamento del perielio di Mercurio e quello, di ben maggiore grandezza, dell'orbita ellittica, in una stella doppia, di una delle due stelle intorno all'altra, causato dalla curvatura dello spazio-tempo; la deflessione della luce per azione gravitazionale e lo spostamento corrispondente delle righe spettrali. Si deve ancora aggiungere che, secondo la teoria einsteiniana, perturbazioni del campo gravitazionale si propagano con onde (onde gravitazionali), e che la velocità con la quale dette perturbazioni si propagano è uguale alla velocità della luce. È cosí data una risposta suggestiva al quesito relativo al modo col quale le azioni gravitazionali si risentono a distanza, togliendo l'assurdo dell'istantaneità dell'azione a distanza. Né deve essere taciuto a questo riguardo che proprio recentemente astronomi dell'Università del Massachusetts sembra abbiano ottenuto una conferma indiretta dell'esistenza di dette onde [cfr. Taylor, Fowler e McCulloch 1979].

1.3. Da Einstein a Planck, Bohr, Broglie, Heisenberg, Dirac, Born.

La teoria della relatività ristretta, a parte la revisione di concetti primari della meccanica newtoniana, indica come condizione limite dell'applicabilità di questa, che tutte le velocità che intervengono nel fenomeno in studio siano *piccole* rispetto alla velocità della luce; la teoria della relatività generale, a parte il diverso significato e la diversa interpretazione che dà a grandezze fondamentali della meccanica di Newton e la diversa concezione che la informa, ne limita l'applicabilità in presenza di fenomeni gravitazionali in scala astronomica.

Esiste però anche una terza limitazione alla meccanica newtoniana, che deriva invece dalla considerazione di moti in scala microscopica, piú precisamente dallo studio del moto delle molecole, degli atomi, dei nuclei, degli elettroni, ossia di piccoli sistemi costituiti da un piccolo numero di particelle elementari; e come le teorie della relatività ristretta e generale di Einstein hanno portato a un riesame profondo dei concetti di spazio, tempo, massa e forza, cosí lo studio dei moti sopra indicati, che è fatto dalla cosiddetta «meccanica quantistica», ha portato a una profonda revisione dei concetti fondamentali di causalità e di determinismo, e pertanto, sotto questo punto di vista è ancora piú rivoluzionario delle stesse teorie di Einstein.

L'iniziatore di questa rivoluzione fu Max Planck, che il 14 dicembre 1900 presentò, in una riunione della Deutsche Physikalische Gesellschaft a Berlino, una deduzione della legge della radiazione del corpo nero, nella quale introdusse un'ipotesi fondamentale: un oscillatore di frequenza propria ν può assorbire o cedere energia solo in *pacchetti* di valore

$$E_\nu = h\nu \tag{20}$$

dove h è una nuova costante fondamentale della natura, che ha le dimensioni di un'azione (ML^2T^{-1}), ed è nota come «costante di Planck», o anche «quanto di azione», mentre al secondo membro della (20) si dà il nome di «quanto d'energia». L'ipotesi di Planck [1901] porta quindi alla concezione davvero rivoluzionaria (perché in netto contrasto con quella della meccanica classica, nella quale l'energia è una variabile dinamica *continua*), che non soltanto la materia e la carica elettrica posseggono struttura granulare, ma anche l'energia di radiazione possiede la stessa struttura, si compone cioè di «quanti». A porre in tutta evidenza questa proprietà fu ancora Einstein [1905*b*] con le sue ricerche sull'effetto fotoelettrico: all'inizio del secolo era noto come risultato sperimentale che, quando la luce incide sulla superficie di un metallo, da questo vengono espulsi elettroni. L'interpretazione del fenomeno data dalla fisica classica era che l'energia luminosa incidente, di natura elettromagnetica, fosse in parte utilizzata per compiere il lavoro necessario per vincere l'energia di legame dell'elettrone e in parte per comunicargli energia cinetica. Ma in questa interpretazione interveniva soltanto l'intensità della radiazione incidente, non la sua frequenza, e questo era in netto contrasto con il risultato sperimentale che, al di sotto di una certa frequenza, qualunque sia detta intensità, l'effetto fotoelettrico non si produce (esperimento di Lenard), mentre l'energia cinetica comunicata agli elettroni emessi dipende esclusivamente dalla frequenza e cresce linearmente con questa (esperimento di Millikan). Questo aspetto, inesplicabile dal punto di vista dell'ottica ondulatoria, trovò una brillante spiegazione quantitativa con l'ipotesi corpuscolare e quantistica di Einstein, per la quale l'energia in un fascio di luce monocromatica si propaga in pacchetti (o fotoni) di valore $h\nu$; è questo quanto di energia che può essere trasferito a un elettrone, ed è l'eccesso di questo quanto rispetto all'energia di estrazione (*work-function*) W, che corrisponde all'energia cinetica E_e dell'elettrone espulso. Si ha cosí

$$h\nu = W + \frac{1}{2} m_e u^2 \tag{21}$$

se m_e è la massa dell'elettrone e u la sua velocità: la (21) è la famosa equazione fotoelettrica di Einstein, che spiega quegli aspetti dell'effetto fotoelettrico noti a quel tempo.

Successivamente il quanto di azione fu introdotto nei modelli atomici, ad opera specialmente di Niels Bohr, il quale pubblicò la sua teoria atomica nel 1913, sviluppando una teoria, in seguito generalizzata da Sommerfeld e Wilson, che ebbe un grandissimo successo e nella quale il carattere corpuscolare era di nuovo attribuito alla radiazione, ciascun fotone essendo associato a una energia $h\nu$, una massa $h\nu/c^2$, in accordo con la legge di Einstein (19), e una quantità di moto $h\nu/c$. D'altra parte, i fenomeni d'interferenza, di diffrazione e di rifrazione stavano sempre a dimostrare che almeno in un grande numero di problemi la teoria ondulatoria della luce conservava la sua piena validità; di qui, il dualismo, per il quale per spiegare i risultati sperimentali era necessario ammettere che la luce avesse a un tempo natura ondulatoria e natura corpuscolare.

A complicare le cose venne poco piú tardi la scoperta che uguale comportamento hanno anche le particelle *materiali*, cioè le particelle che hanno una massa a riposo m_0 diversa da zero, quali gli elettroni, i protoni, i neutroni, i mesoni, le molecole, ecc. Nel caso dei fotoni, le proprietà ondulatorie furono scoperte per prime e le proprietà corpuscolari in seguito; nel caso delle particelle materiali, le scoperte avvennero in ordine inverso e, a causa di questo ordine di successione storica, era invalsa la credenza che la luce fosse costituita da onde, e che ad esempio gli elettroni fossero corpuscoli, ma nella realtà è da credere che *tutte* le particelle, fotoni, elettroni, ecc. siano simili, nel senso che tutte possiedono *alcune* proprietà corpuscolari, e *alcune* proprietà ondulatorie.

Per spiegare questo dualismo, e per spiegare inoltre le numerose proprietà dei modelli atomici che le ricerche teoriche e sperimentali avevano messo in evidenza, furono ideate a partire dal 1925 *tre* meccaniche quantistiche – rispettivamente da Louis de Broglie e Erwin Schrödinger; Max Born, Werner Heisenberg e Pascual Jordan; Paul-Adrian-Maurice Dirac – che nella realtà, sia pure usando algoritmi diversi, ne costituiscono una sola. Qui ci si limita a fare cenno alla meccanica di Broglie e Schrödinger, che è chiamata, a segnalarne l'aspetto piú caratteristico, «meccanica ondulatoria» e che segue nei suoi procedimenti da vicino l'analogia tra i principî della meccanica e quelli dell'ottica [i metodi e gli sviluppi di essa sono mirabilmente esposti in Wichmann 1967].

Il principio ispiratore dell'opera di Louis de Broglie [1923] fu che, come l'ottica geometrica richiede un'estensione se si vuole considerare il campo dell'ottica fisica, cosí la meccanica deve essere estesa se con essa si vogliono studiare i fenomeni quantici; e poiché la base dell'ottica fisica è l'equazione delle onde, il problema da risolvere per detta estensione era di scoprire l'equazione delle onde nella meccanica quantistica. Si supponga a questo scopo che a ogni particella in moto sia associata un'onda, e che la particella si muova in assenza di ogni forza esterna; siano E l'energia, $\mathbf{p}$ la quantità di moto ($\mathbf{p}=m\mathbf{u}$), e m la massa della particella. Sembra logico ammettere che l'onda ad essa associata si muova nella sua stessa direzione, e la si possa rappresentare con la funzione d'onda complessa

$$\psi\left(\mathbf{x}, \frac{\mathbf{p}}{\hbar}; t\right) = A \exp(i\mathbf{x}\cdot\mathbf{k} - i\omega t), \tag{22}$$

dove i è l'unità immaginaria ($i^2=-1$); $\hbar=h/2\pi$; A è l'ampiezza costante dell'onda, $\mathbf{k}$ è il vettore d'onda e ω la frequenza.

Il primo problema è di trovare la relazione tra $\mathbf{k}$ e ω, che descrivono l'onda, e le grandezze E, $\mathbf{p}$, m che descrivono la particella. Ricercando come si trasformano $\mathbf{k}$ e ω rispetto alla trasformazione di Lorentz, e ponendo la condizione dell'invarianza relativistica rispetto a detta trasformazione, si arriva al seguente risultato

$$E=\hbar\omega \qquad \mathbf{p}=\hbar\mathbf{k} \tag{23}$$

e pertanto, se λ è la lunghezza d'onda corrispondente a $\mathbf{k}$, $\lambda=h/p=2\pi/k$. La seconda delle (23) è chiamata «equazione di Broglie», mentre λ è chiamata la lunghezza d'onda di Broglie della particella, e la ψ data dalla (22) funzione d'onda di Broglie. Si noti peraltro che, in realtà, le relazioni tra energia, quantità di moto, frequenza e vettore d'onda, che risultano dalla condizione di soddisfare il principio di relatività ristretta, sono $E=C\omega$, $\mathbf{p}=C\mathbf{k}$, dove C è una costante caratteristica della particella, e non c'è motivo per cui C debba essere la *stessa* per tutte le particelle, e quindi uguale a quella corrispondente ai fotoni. Il prendere $C=\hbar$ è una pura ipotesi, peraltro pienamente confermata dai risultati sperimentali.

Si può ora riconoscere il perché le onde materiali non abbiano considerazione nella meccanica macroscopica: in effetti, per $m=10^{-5}$ g e $u=1$ cm/s, poiché $h=6{,}626\times10^{-27}$ erg sec., usando l'espressione della lunghezza d'onda di Broglie si ottiene $\lambda\simeq6{,}6\times10^{-22}$ cm, davvero troppo piccola per potere influenzare la meccanica dei corpi ordinari.

Il secondo stadio nella fondazione della meccanica ondulatoria è costituito dalla ricerca dell'equazione alla quale la funzione d'onda soddisfa: l'ipotesi piú importante è che l'equazione d'onda, che descrive una *singola* particella, sia un'equazione differenziale *lineare*, e questo allo scopo di permettere che le soluzioni dell'equazione stessa soddisfino al principio di sovrapposizione, ossia ogni combinazione lineare di dette soluzioni sia essa stessa una soluzione dell'equazione. Ora, l'equazione differenziale piú semplice che sia soddisfatta dalle onde di Broglie è la cosiddetta equazione di Klein-Gordon, che descrive la propagazione dell'onda associata a una particella solo in regioni *vuote* dello spazio-tempo, ossia a grande distanza da tutte le altre particelle, ed è data dalla

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\psi(\mathbf{x},t)-\Delta_2\psi(\mathbf{x},t)=-\left(\frac{mc}{\hbar}\right)^2\psi(\mathbf{x},t) \qquad \Delta_2=\Sigma_j\frac{\partial^2}{\partial x_j^2}. \tag{24}$$

La piú generale sovrapposizione (continua) di onde piane soddisfacenti alla (24) si può rappresentare con la

$$\psi(\mathbf{x},t)=\int_{(\infty)} d^3\left(\frac{\mathbf{p}}{\hbar}\right)A\left(\frac{\mathbf{p}}{\hbar}\right)\exp\left(i\mathbf{x}\cdot\frac{\mathbf{p}}{\hbar}-i\omega t\right) \tag{25}$$

in cui qui $A(\mathbf{k})$ è una funzione complessa del vettore d'onda $\mathbf{k}$, da determinarsi in modo che la $\psi(\mathbf{x},t)$ nell'istante $t=0$, o in un altro istante arbitrario fissato,

sia uguale a una data funzione; mentre l'integrale è esteso all'intero spazio tridimensionale **k**. La ω poi è la funzione di p, che si ricava dalle relazioni

$$E=\hbar\omega=\frac{mc^2}{\sqrt{1-\frac{u^2}{c^2}}} \qquad p=\frac{mu}{\sqrt{1-\frac{u^2}{c^2}}}=\hbar k,$$

e pertanto

$$\omega=\left(\frac{c}{\hbar}\right)\sqrt{p^2+m^2c^2}. \tag{26}$$

È da ricordare infine che, secondo la teoria dell'integrale di Fourier, dalla (25) consegue

$$A\left(\frac{\mathbf{p}}{\hbar}\right)=(2\pi)^{-3}\int_{(\infty)} d^3(\mathbf{x})\,\psi(\mathbf{x},0)\exp\left(-i\mathbf{x}\cdot\frac{\mathbf{p}}{\hbar}\right). \tag{25'}$$

Ovviamente, nella (24) la $A(\mathbf{k})$ deve essere una funzione *ragionevole* di **k**, affinché l'integrale a secondo membro abbia un senso, cosí come nella (25') la $\psi(\mathbf{x})$ deve essere una funzione *ragionevole* di **x** perché l'integrale a secondo membro converga.

Un terzo problema, infine, fondamentale, alla base della meccanica ondulatoria, è quello di dare un significato fisico alla funzione d'onda $\psi(\mathbf{x},t)$: e questo significato fu ipotizzato per la prima volta da Max Born [1926]: la funzione d'onda $\psi(\mathbf{x},t)$, ossia l'ampiezza dell'onda di Broglie, descrive la distribuzione della probabilità della particella nello spazio e nel tempo, e precisamente, la probabilità di trovare la particella, corrispondente alla $\psi(\mathbf{x},t)$, in una piccola regione di volume $d^3(\mathbf{x})$ contenente il punto **x** è proporzionale a $|\psi(\mathbf{x},t)|^2 d^3(\mathbf{x})$. La densità di probabilità è perciò proporzionale al quadrato del modulo della funzione d'onda.

Dai risultati sopra indicati discendono subito importanti conseguenze.

Si supponga che la funzione $\psi(\mathbf{x},0)$ sia grande solo in un piccolo intorno del punto **x**, e nulla al di fuori di esso, cosí che, per il significato che la funzione stessa ha, la posizione della particella sia *ben* definita; dalla (25') risulta allora che la $A(\mathbf{p}/\hbar)$ ha all'incirca la forma di un'onda piana, e pertanto i suoi valori risultano sparpagliati in una regione molto grande dello spazio **p**, il che vuol dire che la quantità di moto della particella è *mal* definita. Inversamente, se è la funzione $A(\mathbf{p}/\hbar)$ ad essere localizzata *molto bene* nello spazio delle quantità di moto, dalla (25) si ricava che è la funzione d'onda $\psi(\mathbf{x},0)$ ad avere all'incirca la forma di un'onda piana, e pertanto ad avere i suoi valori sparpagliati in una regione molto grande dello spazio **x**, e corrispondentemente è la posizione della particella a risultare *mal* definita. La teoria degli integrali di Fourier dà a queste parole una formulazione precisa, che porta a correlare la concentrazione della funzione $A(\mathbf{p}/\hbar)$ con la concentrazione della funzione $\psi(\mathbf{x},0)$, e il risultato è una relazione d'indeterminazione: la precisione con la quale è definita la posizione è inversamente proporzionale alla precisione con la quale è definita la quantità di moto, e precisamente, se si indica con Δx_α la indeterminazione della

posizione x_α, e con Δp_α la indeterminazione della componente della quantità di moto p_α, è

$$\Delta x_\alpha\cdot\Delta p_\alpha\geq\frac{h}{2\pi} \qquad \alpha=1,2,3. \tag{27}$$

Le (27) costituiscono le relazioni di indeterminazione formulate da Heisenberg [1927]. Una relazione analoga alla (27) vale per l'incertezza della grandezza ΔE della variazione di energia di una particella e l'incertezza della grandezza dell'intervallo di tempo Δt durante il quale la variazione è prodotta:

$$\Delta E\cdot\Delta t\geq\frac{h}{2\pi}. \tag{27'}$$

Non si dà qui la deduzione rigorosa delle (27) e (27'); ci si limita ad osservare che esse sono conseguenze necessarie della descrizione ondulatoria della particella, in quanto da essa deriva che, se la posizione di una particella deve essere definita molto bene, il treno d'onda corrispondente deve essere molto corto, condizione questa incompatibile con quella che il vettore d'onda **k**, e quindi la quantità di moto **p** ad essa correlata per le (23), siano ben definite, perché quest'ultima condizione richiederebbe che il treno d'onda abbia la forma di un'onda sinusoidale su un intervallo che comprende un grande numero di periodi completi.

È chiaro che queste condizioni sono in netta antitesi con la proprietà, senz'altro ammessa nella meccanica macroscopica, che posizione e velocità di ogni data particella in un dato istante t possano essere conosciute, e determinate, con precisione arbitraria; pertanto le relazioni (27) e (27') stabiliscono i limiti oltre i quali i concetti della fisica classica non possono piú essere applicati. Cosí ad esempio, se si considerano elettroni mobili, che attraversano una fenditura di larghezza d, e λ è la lunghezza d'onda di Broglie, supposta uguale per tutti gli elettroni, perché questi possano essere considerati come corpuscoli deve essere $d\gg\lambda$, e poiché $\lambda=h/p$, ne viene $dp\gg h$; se è invece $d\approx\lambda$, e quindi $dp\approx h$, le particelle si comportano come onde, nel senso che si vedranno gli effetti di diffrazione caratteristici di queste. È, di conseguenza, la rappresentazione delle particelle come onde di probabilità, che conferisce alle particelle stesse, a seconda dei casi, proprietà ondulatorie o proprietà corpuscolari.

Si può fare un passo avanti, e considerare il moto di una particella in un campo di forze esterne deducibili da un potenziale, che può descrivere l'effetto sulla particella in esame di tutte le altre particelle presenti nel campo: non si giustifica qui l'ipotesi di poter rappresentare l'interazione fra particelle con l'introduzione di una funzione potenziale, e ci si accontenta di dire che questo, se non è rigoroso, appare ragionevole almeno nell'ambito di una teoria approssimata. Questa teoria è quella dell'equazione di Schrödinger, formata per la prima volta da Erwin Schrödinger [1926] poco dopo l'invenzione della meccanica delle matrici di Heisenberg, e che si basa inoltre su un'altra approssimazione, quella che tutte le velocità siano tanto piccole da poter trascurare tutti gli effetti relativistici.

Un'onda di Schrödinger arbitraria, che ha sempre il significato di un'onda di probabilità, può essere ottenuta, come quella prima considerata, dalla sovrapposizione di onde piane della stessa forma delle onde di Broglie, e soddisfa all'equazione

$$-\frac{\hbar^2}{2m}\Delta_2\psi(\mathbf{x},t)+V(\mathbf{x})\psi(\mathbf{x},t)=i\hbar\frac{\partial}{\partial t}\psi(\mathbf{x},t) \tag{28}$$

in cui V è l'energia potenziale della particella; se E è l'energia di questa è poi

$$\hbar\frac{\partial}{\partial t}\psi(\mathbf{x},t)=E\psi(\mathbf{x},t) \tag{28'}$$

e la (28) si può scrivere nella forma

$$-\frac{\hbar^2}{2m}\Delta_2\psi(\mathbf{x},t)=(E-V)\psi(\mathbf{x},t). \tag{29}$$

Perché una soluzione della (28), o della (29), sia *fisicamente possibile*, essa deve essere a quadrato integrabile, ossia $|\psi|^2$ deve essere integrabile nell'intero spazio. Le (28) e (29) sono appunto le famose equazioni di Schrödinger: la (28) è chiamata equazione di Schrödinger dipendente dal tempo, e la (29) equazione di Schrödinger indipendente dal tempo. La prima vale per *tutte* le onde di Schrödinger, mentre la seconda (per un *dato* valore di E) vale solo per quelle che corrispondono ad una particella di energia E. Le (28) e (29) sono state applicate con grandissimo successo a molti problemi della meccanica atomica e molecolare. Ci si limita qui a citare:

1) la quantizzazione del moto dell'atomo e la dimostrazione dell'esistenza di livelli energetici discreti; in effetti si ricava che le soluzioni della (29), aventi una dipendenza temporale della forma $[\exp(-i\omega t)]$, soddisfacenti alla condizione di regolarità sopra definita, si hanno per i valori di ω formanti un insieme discreto ω_1, ω_2, ... e l'energia dell'ennesimo stato stazionario è data da $E_n=\hbar\omega_n$;
2) l'effetto tunnel, ossia l'attraversamento quantomeccanico di una barriera di potenziale superiore all'energia totale della particella.

Si indicano ancora, nell'ambito della semplice teoria di Schrödinger, gli enti matematici, che nella meccanica quantistica corrispondono alle grandezze che definiscono posizione e quantità di moto nella meccanica classica. Per semplicità, ci si limita al caso unidimensionale: per il significato già attribuito alla funzione d'onda $\psi(\mathbf{x},t)$, e supposto che questa sia normalizzata, in modo che $\int_{-\infty}^{+\infty}dx\,|\psi(x,t)|^2=1$, il valor medio di x e quello di x^2 in un dato istante t sono dati dalle

$$\begin{aligned} M(x) &= \bar{x}=\langle\psi|x|\psi\rangle=\int_{-\infty}^{+\infty}dx\,x|\psi(x)|^2 \\ M(x^2) &= \langle\psi|x^2|\psi\rangle=\int_{-\infty}^{+\infty}dx\,x^2|\psi(x)|^2 \end{aligned} \tag{30}$$

e l'indeterminazione Δx di x, definita dalla $\Delta x=\sqrt{M[(x-\bar{x})^2]}$ risulta di conseguenza data dalla

$$(\Delta x)^2=\int_{-\infty}^{+\infty}dx\,(x-\bar{x})^2|\psi(x)|^2=M(x^2)-[M(x)]^2.$$

In modo analogo si calcola il valor medio di una qualsiasi altra grandezza che dipenda solo da x, ad esempio, il potenziale $V(x)$:

$$M[V(x)]=\langle\psi|V(x)|\psi\rangle=\int_{-\infty}^{+\infty}dx\,V(x)|\psi(x)|^2. \tag{30'}$$

Dagli esempi ora dati appare senz'altro come, nota la funzione d'onda corrispondente a un determinato stato della particella, ne risulta senz'altro il valor medio di ogni grandezza funzione della variabile (posizione quantomeccanica) della particella stessa; ma se ci si domanda quale è il valore numerico locale istantaneo di detta grandezza, si deve rispondere che questo valore numerico non è definito: ossia una variabile quantomeccanica *non ha* un valore numerico; essa è definita soltanto con un procedimento per mezzo del quale si può calcolare il suo valor medio per ogni data funzione d'onda (cioè per ogni dato stato della particella).

Per quanto si riferisce alla variabile quantomeccanica quantità di moto, la determinazione del suo valor medio richiede anzitutto la definizione di detta variabile: si consideri a questo scopo una funzione d'onda, che in un intervallo molto ampio ha la forma

$$\psi(x)=A\exp(ix/\lambda)=A\exp\left(\frac{ix\bar{p}'}{\hbar}\right)$$

mentre esternamente a detto intervallo è $\psi=0$. Per quest'onda è certo il valor medio della quantità di moto prossimo a $\bar{p}'$, e di conseguenza si può scrivere

$$\bar{p}'\psi(x)=-i\hbar\frac{\partial}{\partial x}\psi(x);$$

e poiché,

$$\int_{-\infty}^{+\infty}\psi\psi^*(x)\,dx=1$$

se $\psi^*(x)$ è il complesso coniugato di $\psi(x)$, è

$$\bar{p}'\simeq\int_{-\infty}^{+\infty}dx\,\psi^*(x)\left[-i\hbar\frac{\partial}{\partial x}\right]\psi(x).$$

Questa espressione è generalizzata per definire il valore medio della quantità di moto quantomeccanica, assumendo come postulato

$$M(p)=\langle\psi|p|\psi\rangle=\int_{-\infty}^{+\infty}dx\,\psi^*(x)\left[-i\hbar\frac{\partial}{\partial x}\right]\psi(x) \tag{31}$$

per ogni funzione d'onda di Schrödinger. Appare pertanto che la (31) si può interpretare nel senso che nella teoria di Schrödinger la variabile quantità di moto è rappresentata dall'operatore differenziale $p = -i\hbar(\partial/\partial x)$, che agisce sulla funzione d'onda alla sua destra. Poiché il quadrato di un operatore si ottiene operando due volte con lo stesso operatore, è

$$(32) \qquad p^2 = -i\frac{\hbar\partial}{\partial x}\left(-\frac{i\hbar\partial}{\partial x}\right) = -\frac{\hbar^2\partial^2}{\partial x^2}$$

e il valor medio del quadrato della quantità di moto è dato dalla

$$(33) \qquad M(p^2) = \langle\psi^*|p^2|\psi\rangle = \int_{-\infty}^{+\infty} dx\, \psi^*(x)\left[-\frac{\hbar^2\partial^2}{\partial x^2}\right]\psi$$

mentre è $\Delta p = \sqrt{M(p-\bar{p})^2}$. Analogamente, per il valor medio quantomeccanico della funzione hamiltoniana H, che nella meccanica classica dà l'energia totale $(p^2/2m)+V$, si ha

$$(34) \qquad M(H) = \langle\psi^*|H|\psi\rangle = \int_{-\infty}^{+\infty} dx\, \psi^*(x)\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+V\right]\psi(x),$$

di guisa che l'operatore energia H è dato dall'operatore differenziale

$$H = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+V(x).$$

Quanto ora indicato è importante, non soltanto perché fa vedere che mentre non ha senso in meccanica quantistica parlare di valori locali istantanei di una qualsiasi grandezza quantomeccanica, è invece ben definito, per ogni funzione d'onda, il valor medio della grandezza stessa, ma anche perché permette di mettere in relazione la meccanica quantistica con la meccanica classica. In effetti, questa relazione è stabilita dal teorema di Ehrenfest, per il quale i valori medi delle variabili quantomeccaniche soddisfano alle stesse equazioni del moto che sono soddisfatte dalle corrispondenti variabili classiche nella meccanica classica. Cosí è

$$\frac{d}{dt}M(x) = \frac{1}{m}M(p) \qquad \frac{d}{dt}M(p) = -M\left[\frac{dV(x)}{dx}\right]$$

se la funzione d'onda $\psi(\mathbf{x},t)$, che permette di calcolare i valori medi, soddisfa all'equazione di Schrödinger

$$(28) \qquad H\psi(\mathbf{x},t) = i\hbar\frac{\partial\psi}{\partial t}$$

essendo H l'operatore differenziale sopra definito.

Appare pertanto dal teorema ora indicato che la meccanica classica può essere considerata come un caso limite della meccanica quantistica, caso limite corrispondente a ritenere trascurabile ai fini dello studio di un dato fenomeno di moto l'indeterminazione delle variabili, ossia la fluttuazione statistica delle variabili, che è tipica della meccanica quantistica; e la condizione che deve essere soddisfatta perché questo possa essere fatto è quella che già è stata precedentemente discussa, in riguardo alle relazioni di indeterminazione.

La teoria di Schrödinger è stata estesa al caso di sistemi composti non da una sola particella, ma da due o piú particelle interagenti (ad esempio da due elettroni, o un elettrone e un fotone), o anche considerando gli effetti relativistici, e a questo riguardo è fondamentale l'opera di Dirac [1930], al quale si deve il primo trattato completo di meccanica quantistica. Non è possibile qui indicare nemmeno alcuni dei piú grandi progressi della fisica che la meccanica quantistica ha permesso di ottenere, della moltitudine di fenomeni noti da tempo ma rimasti misteriosi e che essa ha interpretato e delle importantissime applicazioni tecnologiche cui essa ha dato luogo. Se questo non può non essere riconosciuto, è tuttavia difficile ammettere che in questo modo si è riusciti ad avvicinarsi a *capire* la natura; se ci si domanda perché non si possa nemmeno ipotizzare di conoscere la precisa posizione di una particella in un dato istante, ma bisogna accontentarsi di conoscere la probabilità di un determinato stato di essa, la risposta potrebbe essere che sarebbe anzitutto necessario definire che cosa s'intende per particella; se questa è un *concentrato* di energia, sembrerebbe difficile definire i confini entro i quali detto concentrato è compreso, ma sarebbe pur sempre necessario dire perché detta probabilità debba propagarsi con onde, e in effetti quanto è stato posto ed accettato ha la sua giustificazione soltanto nel fatto che le conseguenze che ne derivano sono in accordo con i risultati sperimentali e non in un puro ragionamento. Non si può a questo riguardo non accennare a quanto scrisse Einstein a proposito dell'introduzione in modo essenziale del concetto di probabilità: in una lettera da lui inviata a Max Born (datata 3 dicembre 1947) egli dice: «Vedo naturalmente che l'interpretazione di principio statistica... possiede un considerevole contenuto di verità. Tuttavia non posso seriamente credere in essa, perché la teoria contraddice al principio che la fisica deve rappresentare una realtà nello spazio e nel tempo senza fantomatiche azioni a distanza... Io credo fermamente che si perverrà alla fine a una teoria nella quale gli oggetti connessi da leggi non siano delle probabilità ma dei fatti pensati, come si riteneva evidente fino a qualche tempo fa. Non posso tuttavia fornire argomenti logici a sostegno della mia convinzione, ma soltanto chiamare a testimonio il mio dito mignolo, cioè un'autorità che non può pretendere di essere rispettata fuorché dalla mia pelle» [citata in Born 1949, trad. it. p. 153]. E lo stesso Dirac, nella celebrazione del centenario di Einstein all'Unesco, tenuta a Parigi nel 1979, esortò a riconsiderare di nuovo le obiezioni di Einstein alla meccanica quantistica.

È stata molto discussa l'influenza, sul pensiero filosofico e sull'etica, del dilemma *determinismo-indeterminismo* nella fisica. Max Born [1949] riferisce a questo riguardo le parole dette da Ernst Cassirer nel suo trattato *Determinismo e indeterminismo nella fisica moderna*: «Da questo significato della libertà come possibilità circoscritta da leggi di natura nessuna via conduce a quella "realtà" del volere e della decisione pratica con cui ha a che fare la problematica etica. Confondere la "selezione" che secondo la teoria di Bohr l'elettrone può fare tra differenti orbite quantizzate, con una "scelta" nel senso etico di questo

concetto, significherebbe cadere vittime di un equivoco d'ordine puramente linguistico. Una "scelta" infatti c'è solo dove non soltanto sussistono in sé possibilità diverse ma fra di esse si compie una distinzione consapevole e si viene a una decisione consapevole. Proiettare atti del genere nell'elettrone sarebbe una madornale ricaduta in una forma di... antropomorfismo» [Cassirer 1937, trad. it. p. 306]. E piú avanti: «Allora pensare la causalità dominante nella natura, sotto la specie di leggi strettamente "dinamiche" o sotto quella di leggi puramente "statistiche", non cambia nulla... Né sull'una né sull'altra via resta aperto un adito alla sfera della "libertà" che l'etica rivendica per sé» [*ibid.*, p. 308; cfr. anche Born 1949, trad. it. pp. 254-55].

Rimane sempre valida la domanda fino a qual punto lo studio del concetto di moto, quale si è andato evolvendo nei secoli, ha portato a una migliore conoscenza della natura; si crede utile a questo riguardo riportare quanto dice Heisenberg paragonando quanto scrisse Platone nel *Timeo* sulla cosmogonia con quanto è scritto nel trattato di astronomia di Newcomb ed Engelmann: «La descrizione moderna si distingue da quella antica per tre tratti caratteristici: perché pone enunciati quantitativi al posto di enunciati qualitativi; perché riconduce fenomeni differenti alla medesima origine, e perché rinuncia a chiedersi il "perché"» [1932, trad. it. p. 13].

2. *Il moto come è realizzato dagli organismi viventi sulla Terra.*

Dopo l'analisi del moto, quale è concepito dall'uomo, delle sue leggi e delle sue cause, dell'evoluzione dei concetti fondamentali con esso connessi nel tentativo di comprendere il *perché* dei fenomeni di moto, è naturale domandarsi *come* il moto è prodotto, almeno sulla Terra, dove noi viviamo e abbiamo la possibilità di osservare e sperimentare. Tutto in natura si muove, e il movimento è la prima indicazione della vita; le modalità con cui esso avviene sono le piú svariate, ed analizzarle tutte sarebbe impossibile. Qui pertanto ci si limita a considerare casi tipici della locomozione quale è realizzata nel mondo animale, e piú particolarmente della locomozione in acqua, perché le modalità con le quali essa è effettuata sono, forse, piú numerose e piú ricche di insegnamenti che non quelle relative alle locomozioni aerea e terrestre: anche questo studio presenta problemi di enorme difficoltà, che richiedono competenze specifiche in vari campi della scienza, e qui ci si limita ovviamente a considerare solo la parte meccanica dei detti problemi; d'altra parte, questi presentano pure aspetti meravigliosi, che lasciano stupiti per la genialità della soluzione che naturalmente essi hanno avuto.

2.1. Locomozione acquatica.

Tutti gli esseri viventi in acqua traggono la forza di propulsione, grazie alla quale essi si possono muovere a loro piacimento, dalla capacità che essi hanno di imprimere al fluido una variazione di quantità di moto in direzione opposta

a quella del movimento che intendono compiere. Le modalità con le quali essi esplicano questa capacità dipendono dalle dimensioni dell'organismo, dalla forma di questo, da una molteplicità di altre condizioni: le piú frequenti di tali modalità sono quelle qui appresso indicate.

2.2. Propulsione con flagelli (propulsione flagellare).

Microrganismi, come batteri flagellati (ad esempio *Escherichia coli*, *Bacillus subtilis*, *Spirillum*, *Spirochaeta*, *Leptospira*), spermatozoi, piccoli vermi, si muovono nel liquido in cui vivono grazie all'azione propellente esercitata dal liquido stesso sui filamenti, che sporgono dal corpo cellulare cui sono collegati (per batteri e spermatozoi, nel qual caso detti filamenti sono chiamati «flagelli»), o costituiscono essi stessi il corpo del microrganismo (caso dei vermi), e ai quali è impresso un moto ondulatorio.

Questo moto può essere piano, come nel caso dei piccoli vermi, che contraendo ed estendendo i muscoli producono onde trasversali piane di deformazione, propagantisi nella stessa direzione, e generalmente in senso opposto a quello del moto; oppure, come usualmente avviene per batteri e spermatozoi, il moto ondulatorio può essere tridimensionale: precisamente, l'asse del flagello si avvolge ad elica sopra un cilindro, che per semplicità si può supporre circolare, e ogni suo elemento ruota intorno all'asse di detto cilindro. Con riferimento pertanto a un sistema di assi che si trasla, ed eventualmente ruota, con il corpo del microrganismo in studio, essendo l'asse x diretto ed orientato come la velocità di traslazione di questo, la linea secondo la quale è disposto l'asse di un singolo flagello è rappresentata dalle equazioni, in coordinate cilindriche (fig. 1)

$$(1) \qquad r = h \qquad \vartheta = k(x+ct) = kx + \omega t$$

in cui $k = 2\pi/\lambda$ è il numero d'onda e $c = \omega/k$ è la velocità di fase di propagazione dell'onda lungo l'asse del cilindro di raggio h. Nel caso di moto piano, invece, si ha in un riferimento cartesiano

$$(1') \qquad y = h \sin k(x+ct)$$

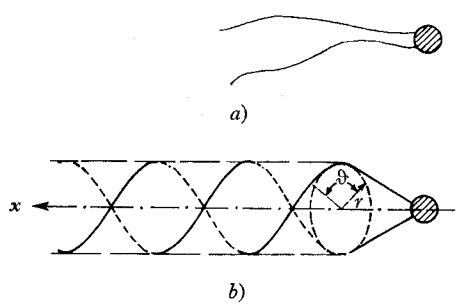


Figura 1.
Propulsione flagellare: *a*) flagelli a riposo; *b*) flagelli in fase di propulsione.

essendo ora h l'ampiezza dell'onda di deformazione, mentre gli altri simboli hanno il significato già dato. In questa seconda eventualità, come la deformazione dell'asse del filamento sia ottenuta è già stato detto sopra; nel primo caso le opinioni sono discordanti.

Secondo alcuni biologi [Berg e Anderson 1973; Silverman e Simon 1974] ciascun flagello è attaccato per mezzo di una specie di giunto universale a una specie di alberino A (fig. 2) che passa attraverso la parete della cellula organica C e termina con un anellino M, libero di ruotare nella membrana citoplasmica; di fronte a M si affaccia un altro anello S che è invece fisso alla parete cellulare. Si suppone che ioni di idrogeno fluiscano nell'interstizio tra A e C, e tra gli anelli M e S, creando un gradiente di potenziale elettrochimico e generando cosí una coppia che fa ruotare A ed il flagello ad esso unito. Secondo altri fisiologi [Calladine 1974], invece, il flagello, che si attacca direttamente alla parete cellulare, si dispone ad elica sul cilindro di raggio h senza ruotare, ossia senza *torcersi*, inflettendosi simultaneamente in due piani normali l'uno all'altro. Tocca ovviamente ai biologi chiarire la questione e risolvere la controversia: qui va solo espressa l'ammirazione per l'organismo unicellulare che ha saputo costruire (non poco prima dell'uomo) un motore elettrico miniaturizzato nel quale vengono utilizzate energia pulita e una materia prima, l'idrogeno, il cui sfruttamento è tuttora un'aspirazione dell'uomo, se è corretta la prima interpretazione; o che ha insegnato come si avvolgono senza torsione i fili in un trefolo e i trefoli in una fune elicoidale, se è esatta la seconda ipotesi.

Le dimensioni dei flagelli, o dei filamenti, sono sempre assai piccole: i singoli flagelli dei batteri possono avere un diametro di circa $0{,}01 \div 0{,}02$ μm (ossia $10 \div 20 \times 10^{-6}$ mm), e lunghezze da 1 a 5 μm. In genere, ogni batterio porta un certo numero di flagelli, che si riuniscono a fasci durante il moto, cosí da avere

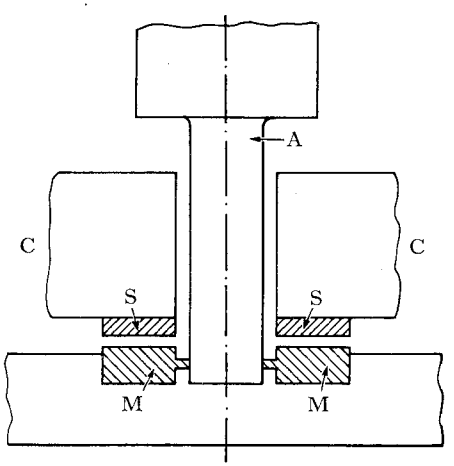


Figura 2.
Propulsione flagellare: sistema di comando dei filamenti secondo Berg e altri.

l'apparenza di un singolo filamento di diametro alquanto maggiore del singolo flagello: l'*Escherichia coli* ad esempio porta sei flagelli ad un estremo, che ad organismo fermo sono disposti con una orientazione arbitraria, mentre nella fase propellente sono riuniti in gruppo; quando questo cambia orientazione, la cellula cambia la direzione del suo movimento. Le cellule eucariotiche hanno flagelli di diametro alquanto maggiore, dell'ordine di 0,2 μm, e di lunghezza fino a $10 \div 100$ μm, mentre per gli spermatozoi i valori corrispondenti di dette grandezze sono $0{,}2 \div 1$, $10 \div 1000$ μm; ed infine, per i piccoli vermi: $20 \div 40$, 1000 μm. In ogni caso il numero di Reynolds Re, ossia il parametro, che è una misura del rapporto tra forza d'inerzia e forza viscosa agente su ogni elemento, e che è dato da $\mathrm{Re} = Ud/\nu$, in cui U è la velocità di traslazione del microrganismo, d il diametro del filamento e ν la viscosità cinematica del fluido in cui il movimento avviene, è sempre molto piccolo variando da 10^{-6}, nel caso dei batteri, a 10^{-3} per gli spermatozoi, a 10^{-1} per i piccoli vermi. Conseguentemente, per lo studio del moto nel campo fluido, e delle forze che questo trasmette al corpo in esso immerso, si possono trascurare le azioni d'inerzia rispetto a quelle viscose, nel qual caso le equazioni che definiscono il movimento stesso sono le equazioni di Stokes linearizzate:

$$-\nabla p + \mu \Delta_2 \mathbf{u} = 0 \tag{2}$$

(in cui p è la pressione, $\mathbf{u}$ la velocità del fluido in un punto generico del campo, μ la viscosità dinamica, mentre ∇ e Δ_2 stanno ad indicare rispettivamente il gradiente e il laplaciano $\Sigma_i\, \partial^2/\partial x_i^2$). L'integrazione della (2), in modo da soddisfare alle condizioni al contorno esprimenti che la velocità del fluido relativa al corpo, in ogni punto della superficie di questo, è nulla, si può ottenere distribuendo sulla superficie stessa (o in opportuni punti interni al corpo) opportune singolarità, e precisamente dipoli di pressione (*Stokeslet*), dipoli di vorticità (*Rotlet*) e infine dipoli di velocità (*Doublet*). I particolari del procedimento si possono trovare nelle numerose ricerche fatte al riguardo; qui ci si limiterà a dare alcuni dei risultati piú interessanti, omettendo le formule troppo complicate per trovare posto in questa sede, tratti essenzialmente dalle note di Chwang e Wu [1971] e osservando che la trattazione ora citata è basata sulla seconda delle ipotesi sopra indicate circa il modo col quale il movimento dei flagelli è realizzato.

Nel caso di filamento avvolto ad elica, e onda di deformazione che si propaga lungo l'asse di questa secondo la (1), l'azione dinamica risultante trasmessa dal fluido sul flagello, nell'ipotesi che il numero delle lunghezze d'onda sul filamento sia intero, è costituito da una forza nella direzione dell'asse dell'elica e di senso contrario alla velocità di propagazione dell'onda, che produce la traslazione del microrganismo, e da una coppia intorno all'asse sopraddetto, che ne determina la rotazione intorno a questo in senso contrario al senso apparente di rotazione dell'elica (corrispondente a ω). Poiché l'azione dinamica risultante sul microrganismo si può considerare sempre *nulla*, dato che l'inerzia di questo è del tutto trascurabile, per effetto della traslazione $U\mathbf{i}$, e della rotazione $-\Omega\mathbf{i}$ si devono esercitare sulla testa del microrganismo (convenendo di chiamare testa tutta la

parte di questo che non è flagello, e che *porta* i flagelli stessi) una forza e una coppia uguali e contrarie a quelle applicate sui flagelli, ed è ponendo queste due condizioni che si ricavano le equazioni determinatrici del moto del microrganismo. È interessante osservare che se in queste si fa tendere a zero il raggio della testa (caso dei piccoli vermi) si ricava che la velocità di traslazione diventa nulla, il che invece non accade nel caso che il moto di deformazione sia un'onda piana (che non dà luogo a nessuna coppia), caratterizzata dalla (1'): risulta cosí evidente un'ottima ragione per la quale i microrganismi senza testa, come i piccoli vermi, adottino per propellersi il moto ondulatorio piano, invece che quello elicoidale.

Si noti ancora che nell'altro caso limite, in cui il rapporto a/b tra il raggio della testa (supposta sferica) e il raggio del flagello (supposto cilindrico) tende ad infinito, è di nuovo $U \to 0$, ossia se la testa è troppo grande, i flagelli non sono sufficienti a spingerla, e questo dà un'ottima ragione per la quale il sistema di propulsione a flagelli non sia usato che da organismi filiformi dalle dimensioni come quelle sopra indicate; e d'altra parte, per organismi cosí piccoli, detto sistema è forse anche l'unico che può essere usato con rendimento non eccessivamente piccolo. L'essere $U = 0$ ad entrambi gli estremi $a/b \to 0$, e $a/b \to \infty$, fa presumere che, rimanendo costanti gli altri parametri da cui U dipende, esista un valore *ottimo* di a/b, per il quale U risulti massima, ed in effetti appare che tale valore massimo è compreso tra 15 e 40, per un'inclinazione β dell'elica prossima a 45°. Ora è notevole il fatto che i raggi dei fasci dei flagelli, per almeno parecchi dei batteri, siano tali che il rapporto a/b corrispondente cada nell'intervallo ora detto, di guisa che è legittimo ritenere che l'aggregazione di un certo numero di flagelli, della quale si è già fatto cenno sopra, nella fase propellente sia prodotta allo scopo di realizzare la massima efficienza di propulsione. Per le cellule eucariotiche, invece, per le quali il raggio della testa è prossimo a 20 μm, il rapporto a/b sembra alquanto piú elevato di quello ottimo; si deve peraltro osservare che il flagello ad esempio della *Euglena viridis* è avvolto da ciglia, che ne producono un aumento effettivo, e forse anche, durante il moto, virtuale del diametro, cosí da riportare il valore di a/b piú vicino a quello ottimo.

Interessa ovviamente anche conoscere il rendimento η del sistema propulsivo, definito come rapporto tra la potenza necessaria per propellere la sola testa e la potenza totale spesa per produrre il moto, e appare che tale rendimento ha il suo valore massimo per un'inclinazione β dell'avvolgimento ad elica dell'asse del flagello prossimo a 45°, e per questo valore di β è molto prossimamente $\eta = \Omega/\omega$: ora, non è inutile osservare che il valore ottimo di β ora indicato è pure quello che corrisponde al massimo rendimento dei meccanismi che utilizzano la coppia elicoidale (ad esempio vite senza fine), mentre, ponendo, in condizioni prossime al massimo rendimento, $\Omega = \omega - \varepsilon$ $(\varepsilon/\omega \ll 1)$ e perciò $\eta = 1 - \varepsilon/\omega$, si ha per η un'espressione che coincide con l'espressione (pure approssimata) del rendimento del sistema propellente indicato in aeronautica come «elica intubata».

Naturalmente, per poter calcolare la massima velocità che un organismo può raggiungere, sarebbe necessario conoscere la massima potenza che esso può sviluppare; purtroppo, non molto è noto riguardo a questo. Solo si può affermare, con sufficiente esattezza, che essa è proporzionale alla lunghezza dei flagelli, e pertanto cresce con la massa in misura alquanto inferiore a quanto avviene, in generale, per i macrorganismi. Comunque, valori di velocità massima sono stati sperimentalmente osservati, ed essi appaiono poco dipendenti dalla massa del microrganismo, e sono dell'ordine di $50 \div 60$ μm/s per i batteri; $70 \div 80$ μm/s per gli spermatozoi; 500 μm/s per i vermi. Per i valori della lunghezza ℓ dei flagelli già indicati, le velocità ora date variano da 20 ℓ/s per i batteri a 0,5 ℓ/s per i piccoli vermi (e per gli spermatozoi di maggiori dimensioni); appare pertanto che, alla scala batterica, i batteri sono dei veri velocisti (la loro velocità alla scala umana corrispondendo a circa 120 km/h). Quanto alla velocità angolare con la quale i flagelli sembrano ruotare intorno all'asse dell'elica, essa appare essere pure ragguardevole, dell'ordine di 100 giri/s, mentre quella della testa è alquanto minore.

Interessa ancora stabilire un confronto tra il sistema propulsore basato sull'onda elicoidale e quello sull'onda piana: da esso si ricava che il primo può essere considerato nettamente superiore al secondo per quanto si riferisce alla velocità di traslazione che può imprimere e per quanto riguarda il rendimento propulsivo per $a/b > 10$, mentre l'opposto avviene per $a/b < 5$, e questi valori sono in buon accordo con la scelta fatta dai microrganismi, le cui caratteristiche geometriche corrispondono alle disuguaglianze sopraindicate.

Ci sono ancora altre proprietà che è interessante indicare:

1) È stato sperimentalmente osservato [Berg e Turner 1979] che parecchie specie di batteri nuotano piú rapidamente in soluzioni diluite di sostanze viscose (almeno fino a valori della viscosità di $2 \div 5$ cP) di quanto non facciano nel solvente puro di minore viscosità: questo avviene ad esempio per l'*Escherichia coli*, e in misura notevolmente maggiore per la *Leptospira*, la cui configurazione geometrica peraltro è notevolmente differente da quella dell'*Escherichia*, e per la quale il massimo della velocità di traslazione si raggiunge per valori della viscosità di circa 300 cP. Questo comportamento si verifica nel caso di soluto caratterizzato da una lunga catena molecolare lineare non ramificata, come la metilcellulosa, e sembra attribuibile alla strutturazione che il mezzo fluido assume nelle condizioni indicate; precisamente, si ritiene che il soluto formi una rete quasi rigida, le cui maglie possono facilmente essere attraversate da particelle di dimensioni non superiori a un dato limite, dipendente dalla concentrazione della soluzione. Il mezzo cosí strutturato si comporta come se avesse due coefficienti di viscosità, l'uno macroscopico elevato, l'altro microscopico alquanto minore, corrispondente alla viscosità del solvente: i flagelli, che operano sul mezzo piú viscoso, trovano in questo un miglior *appoggio* per produrre la forza di trazione, mentre la testa, che avanza nel mezzo meno viscoso, non trova in questo una maggiore resistenza al suo moto.

2) Parecchie cellule eucariotiche, se poste in mezzi di differente viscosità, mentre mantengono invariato il rapporto (ampiezza/lunghezza d'onda), che è generalmente quello corrispondente alla massima efficienza per la propulsione, variano gli altri parametri che caratterizzano la forma d'onda, dimostrando cosí

una loro capacità di adattamento al mezzo in cui si muovono; questo mostra altresí, pure in organismi cosí semplici, l'esistenza di un sistema di controllo dotato di retroazione (*feedback*) che agisce sul sistema di comando degli organi motori [Holwill 1977].

3) Come già è stato detto, il senso della traslazione del microrganismo, che compie un moto ondulatorio, è generalmente contrario al senso di propagazione dell'onda di deformazione da esso prodotta; dal punto di vista meccanico, questo appare conseguenza della seguente proprietà: se si considera un elemento d'asse ds del flagello, e si indicano con $\mathbf{u}_s$ e con $\mathbf{u}_n$ i componenti delle velocità di detto elemento relative al fluido rispettivamente nella direzione ad esso tangente e nella direzione normale, i corrispondenti componenti della forza trasmessa dal fluido all'elemento si possono esprimere con le $d\mathbf{F}_s = -c_s\mathbf{u}_s ds$ e $d\mathbf{F}_n = -c_n\mathbf{u}_n ds$, essendo c_s e c_n i cosiddetti «coefficienti di forza tangenziale e normale», che la teoria in molti casi riesce a determinare. Ora, generalmente, il coefficiente di forza normale c_n è all'incirca doppio, o almeno alquanto piú grande, del coefficiente di forza tangenziale, ed è per questa proprietà che i due sensi di traslazione dell'organismo e di propagazione dell'onda sono opposti. Vi sono casi, però, nei quali, in conseguenza di particolarità strutturali degli organi propellenti, la proprietà sopraddetta si inverte: cosí ad esempio nella *Ochromonas* ogni flagello porta due serie di ciglia (o mastigonemi), una sopra ciascun lato del filamento, che sporgono per circa 1 μm da questo. In queste condizioni il rapporto c_s/c_n risulta maggiore di uno, invece che minore, e conseguentemente il senso della traslazione risulta indentico a quello di propagazione dell'onda (fig. 3). La stessa circostanza si verifica per alcuni piccoli vermi, che presentano una

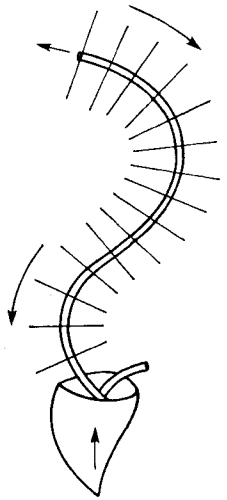


Figura 3.
Flagello con mastigonemi della *Ochromonas*.

superficie caratterizzata da una forte rugosità, ad esempio dovuta a sporgenze anelliformi: questa proprietà è stata per la prima volta messa in luce da Taylor [1952] per il verme marino *Nereis diversicolor*.

2.3. Propulsione con ciglia (propulsione ciliare).

Al crescere delle dimensioni del microrganismo il sistema di propulsione con flagelli si mostra di sempre piú difficile realizzazione per le difficoltà strutturali che l'aumento del numero dei filamenti, indispensabile per ottenere la spinta necessaria, verrebbe a causare. Per superare tali difficoltà sono necessarie profonde modifiche della struttura del sistema propellente, che viene ancora a presentare come elementi attivi per la spinta sottili filamenti, ma questi in numero di gran lunga superiore a quello indicato per le specie considerate precedentemente, e precisamente almeno dell'ordine di qualche centinaio; non sono portati dal microrganismo solo a una estremità di questo, ma sono distribuiti su tutta, o su gran parte della sua superficie, dalla quale sporgono con l'apparenza di piccoli peli, e per questa loro apparenza sono chiamati «ciglia». Il diametro di un ciglio è generalmente di circa 0,25 μm mentre la sua lunghezza varia da 5 a 1200 μm. Le ciglia piú lunghe sono riunite in modo da formare strutture composte, e precisamente, mentre le ciglia singole raramente hanno una lunghezza maggiore di 15 ÷ 20 μm, le ciglia negli aggregati risultanti dall'unione di un numero di esse dell'ordine del migliaio hanno una lunghezza da 25 a 100 μm, e quelle negli aggregati di centinaia di migliaia di componenti sono lunghe fino a 1 mm. Questa tendenza all'aggregazione col crescere della lunghezza delle ciglia è una naturale conseguenza strutturale, in quanto solo in questo modo le ciglia stesse possono acquistare la robustezza necessaria per sopportare la maggiore resistenza al loro moto opposta dal fluido, ma probabilmente è richiesta anche da esigenze energetiche.

La lunghezza delle ciglia poi è generalmente molto minore di quella del corpo ($\ell \ll L$), al contrario di quanto avviene nei microrganismi flagellati, ma per ottenere una propulsione sufficiente il loro numero deve essere molto maggiore di L/ℓ, e pertanto è $n\ell \gg L$. A titolo di esempio si dànno nella tabella 1 gli elementi caratteristici per alcuni dei microrganismi ciliati (fig. 4). Da notare che le

Tabella 1.
Elementi caratteristici relativi a microrganismi ciliati.

Nome	*Gruppo*	*Lunghezza del corpo L (μm)*	*Numero delle ciglia n*	*Lunghezza delle ciglia ℓ (μm)*
Uronema	Protozoo	25	200	5
Paramecium	Protozoo	210	5000	12
Opalina	Protozoo	400	10^5	15
Pleurobrachia	Ctenoforo	15 000	200 (10^8)	500

dimensioni dei piú grandi microrganismi flagellati uguagliano quelle dei piú piccoli microrganismi ciliati, ma i primi si muovono con una velocità dieci volte piú piccola, segno evidente che il loro sistema di propulsione non è piú adatto alle loro dimensioni.

Per produrre il movimento dell'organismo cui appartiene, ognuna delle ciglia appare vibrare; in effetti l'intero ciclo del suo moto oscillatorio consiste di due parti: l'una corrisponde a quella che si può chiamare «corsa di lavoro», durante la quale si mantiene eretta e ruota quasi rigidamente intorno alla sua base di attacco al corpo, mentre l'altra costituisce la «corsa di ritorno» e in essa un'onda di flessione si propaga dalla base alla sommità piegando il filamento e mantenendolo per la piú gran parte in prossimità della superficie, dove la velocità relativa ad esso del fluido è quasi nulla. Mentre durante la corsa di lavoro il moto si può considerare piano, nella corsa di ritorno, in generale, la estremità di ognuna delle ciglia si sposta lateralmente da una parte all'altra dell'attacco di essa alla superficie del corpo.

I movimenti delle varie ciglia sono coordinati, cosí come i remi dei vogatori di un'imbarcazione da regata, ma mentre per questa i remi sono sincronizzati, e pertanto non c'è sfasamento tra i cicli corrispondenti ai vari vogatori, il moto delle ciglia presenta metacronia, ossia quelli che appartengono a una medesima fila vibrano in sincronia, ma sono sfasati rispetto alle ciglia delle file attigue, e questo provoca nell'insieme quel movimento a onde che ricorda l'effetto di un campo di grano percorso dal vento (onde metacronali): la cresta dell'onda corrisponde alle ciglia che sono nella stessa fase del moto oscillatorio, appartenente generalmente alla corsa di lavoro, mentre il cavo dell'onda è corrispondente alle ciglia che si trovano in una stessa altra fase, corrispondente alla corsa di ritorno.

Secondo Knight-Jones [1954] si possono distinguere quattro tipi di metacronismo, a seconda della relazione tra la direzione della propagazione delle onde e

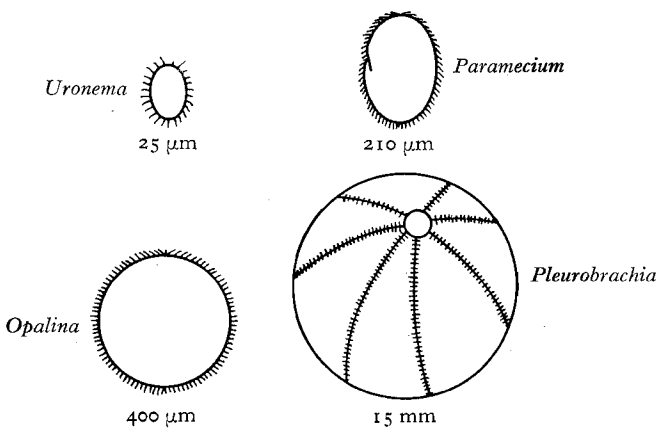


Figura 4.
Microrganismi tipici con propulsione ciliare.

quella della corsa di lavoro delle ciglia: il metacronismo è detto simplettico se le onde si propagano nello stesso senso in cui si muovono le ciglia nella corsa di lavoro, e detto metacronismo dà alla superficie ciliata l'apparenza di un inviluppo continuo ondulato. Nel metacronismo antiplettico le onde si propagano in senso opposto a quello delle ciglia nella corsa di lavoro, e questo fa apparire i fasci di ciglia come ruote a pale. Infine si può avere il metacronismo diaplettico, se i fronti delle onde metacronali non sono trasversali alle fila di ciglia, ma inclinate, di guisa che le onde stesse appaiono come il bordo del filetto di una vite rotante, e a seconda che questa è destra o sinistra le onde sono dette dexioplettiche o laeoplettiche.

Lo studio del sistema di propulsione a ciglia vibratili si può fare col metodo delle singolarità, in modo analogo a quello accennato per il sistema a flagelli. La trattazione piú completa sull'argomento è quella svolta da Blake [1972], e, sebbene parecchi siano ancora i quesiti non sufficientemente chiariti, i risultati ottenuti sono già di grande interesse. Da essi appare che l'effetto propellente delle vibrazioni delle ciglia si esplica non tanto nel raccogliere dal fluido e trasmettere al corpo del microrganismo una forza nel senso di avanzamento cosí da vincere la resistenza al movimento, che per effetto di questo il fluido esercita sulla superficie del corpo stesso, quanto nel ridurre od annullare quasi, in media, detta resistenza: in effetti, la forza media sulle ciglia, durante un ciclo, è circa nulla, mentre dal diagramma della velocità media del fluido relativa al microrganismo appare che la tensione tangenziale alla superficie è quasi nulla. Il sistema di propulsione corrispondente al metacronismo antiplettico, o anche dexioplettico, appare inoltre essere piú efficiente di quello simplettico, probabilmente perché per esso gli effetti d'interferenza tra le ciglia e tra queste ed il corpo del microrganismo sono piú favorevoli. Il confronto tra i risultati teorici e quelli sperimentali non è sempre chiaro, ma non mostra contraddizioni non spiegabili. A titolo di esempio si espongono nella tabella 2 alcuni dei risultati piú significativi.

Tabella 2.
Confronto fra risultati teorici e sperimentali.

	Uronema	*Opalina*	*Paramecium*	*Pleurobrachia*
$R_b = \frac{\rho U L}{\mu}$	10^{-2}	10^{-2}	10^{-1}	10^{3}
$R_c = \frac{\rho \sigma \ell^2}{\mu}$	10^{-3}	10^{-2}	10^{-2}	10
Frequenza delle vibrazioni σ (s^{-1})	?	$1 \div 4$	$10 \div 25$	$5 \div 15$
Velocità adimensionata $U/\sigma L$ (calcolata)	?	$0{,}15 \div 0{,}25$	$0{,}5 \div 1$	$0{,}75 \div 1{,}25$
Velocità calcolata U ($\mu m/s$)	?	$14 \div 94$	$314 \div 1570$	$1{,}4 \div 7{,}1\,(\times 10^4)$
Velocità osservata ($\mu m/s$)	1150	$30 \div 100$	$500 \div 2500$	$1 \div 5\,(\times 10^4)$
Metacronismo	?	Simplettico	Dexioplettico	Antiplettico

Ancora una volta appare che il microrganismo *relativamente* piú veloce è il piú piccolo, mentre risultano chiaramente le scarse qualità natatorie della *Opalina*, le quali si raccordano meglio con quelle corrispondenti ai microrganismi flagellati che non a quelli ciliati. Da osservare d'altra parte che è stato sperimentalmente osservato che il *Paramecium*, che ha metacronismo dexioplettico in condizioni normali, col crescere della viscosità del mezzo tende a cambiare la coordinazione del moto delle ciglia, cosí da realizzare ai piú alti valori della viscosità (dell'ordine di 100 cP) il metacronismo simplettico; le ragioni per cui questo avviene (ragioni energetiche?) non sono chiare, ma ad ogni modo questa circostanza dimostra, in analogia a quanto già è stato osservato per le cellule eucariotiche, una capacità del microrganismo di adattamento al mezzo in cui vive, e quindi l'esistenza di un sistema di controllo atto ad agire e a reagire sul sistema di comando.

2.4. Propulsione anguilliforme.

Come già per il sistema di propulsione a flagelli, cosí anche quello a ciglia vibratili ha una limitazione per quanto riguarda le dimensioni del microrganismo che utilizza detto sistema; in effetti, se k è una opportuna costante adimensionale, dipendente dalla forma del microrganismo, la resistenza che questo presenta in corrente fluida di velocità U, può essere espressa con la $D = k\mu LU$, e pertanto, se P è la potenza spesa per ognuna delle ciglia nella propulsione, il rapporto $\eta = k\mu LU^2/nP$ si può assumere come misura dell'efficienza del sistema propellente, e poiché $U \propto \sigma\ell$, $F \propto \sigma\ell^2\mu$, $P \propto \sigma^2\ell^3\mu$, appare che è $\eta \propto (L/n\ell)$, e potendosi porre il numero n delle ciglia che possono stare sulla superficie proporzionale a questa, e quindi a L^2, si ha che l'efficienza è inversamente proporzionale a L, il che pone ovviamente un limite a L. Questa può non essere l'*unica* ragione, ma certamente è anche *una* ragione, perché i microrganismi ciliati hanno una lunghezza massima dell'ordine del millimetro; l'aggregazione di un numero notevole di ciglia, cosí da formare specie di pale remiganti flessibili, può accrescere il campo di applicazione del sistema di propulsione, ma limitatamente ad organismi di poche decine di millimetri di lunghezza.

Per dimensioni maggiori si hanno cambiamenti strutturali, morfologici, che rendono possibili altri sistemi di propulsione, che per i regimi di deflusso corrispondenti a queste maggiori dimensioni hanno piú efficienza di quelli fin qui considerati. In effetti, con l'aumento delle dimensioni dell'organismo, che generalmente è accompagnato da un aumento della sua velocità, cresce il numero di Reynolds del flusso, e questo viene ad assumere un valore per il quale le forze d'inerzia del fluido non soltanto non possono piú essere trascurate rispetto alle forze viscose, ma sono del tutto preponderanti rispetto a queste, e dal regime di Stokes, quale finora si è considerato, si passa prima a un regime laminare, e successivamente a un regime turbolento di deflusso.

Nelle condizioni ora dette il sistema di propulsione, che è piú largamente usato da un grandissimo numero di specie di animali acquatici è quello detto anguilliforme, perché proprio, fra gli altri, del genere *Anguilla*: esso consiste

in un moto ondulatorio, analogo a quello già indicato per i microrganismi flagellati, con la differenza, rispetto a questo, che è generalmente piano, e non elicoidale, ed inoltre che l'ampiezza dell'onda cresce a mano a mano che questa si propaga verso la coda.

Nel modo anguilliforme puro è tutto il corpo che prende parte al moto ondulatorio, di guisa che se si rappresenta l'asse di detto corpo in due posizioni assunte negli istanti (t) e $(t+\Delta t)$, le corrispondenti configurazioni appaiono essere quelle indicate nella figura 5. Si deve inoltre osservare che le forme dei pesci che usano questo sistema di propulsione (come ad esempio quelli degli Attinotterigi e Allotriognati) sono tali che le loro dimensioni ad angolo retto con la direzione del moto sono molto piccole rispetto alla loro dimensione L, mentre l'area della loro sezione trasversale varia molto lentamente. Nell'ipotesi che anche gli spostamenti prodotti dal moto ondulatorio siano piccoli rispetto a L, si può applicare la cosiddetta teoria dei corpi sottili (*slender body theory*) per determinare il flusso. Se U è la velocità di traslazione del pesce, che si immagina avvenire secondo l'asse di questo nella posizione indeformata (che si assume come asse x), con riferimento a un sistema di assi (x, y, z) che si trasla con la stessa velocità U, sia $h(x,t)$ lo spostamento del baricentro della sezione trasversale del corpo, di ascissa x, nella direzione dell'asse z, prodotto dal moto ondulatorio prima indicato. La *slender body theory* conduce a determinare il flusso in ogni sezione $x = \text{cost.}$ ad ogni dato istante t come quello prodotto dal cilindro, la cui direttrice è uguale al contorno della sezione del corpo con lo stesso piano $S(x)$ e che si trasla nella direzione dell'asse z rispetto al fluido in quiete all'infinito con la velocità $w(x,t) = (\partial h/\partial t) + U(\partial h/\partial x)$ (fig. 6). In queste condizioni, per valori del numero di Reynolds abbastanza elevati come quelli che corrispondono al caso ora in esame, il fluido può essere considerato come *perfetto*, agli effetti del calcolo della forza di propulsione, e questa può essere espressa con la formula ottenuta da Lighthill [1970]

$$\bar{T} = \frac{1}{2}\rho A(L)\left\{\left(\frac{\overline{\partial h}}{\partial t}\right)^2 - U^2\left(\frac{\overline{\partial h}}{\partial x}\right)^2\right\}_L \tag{3}$$

in cui ρ è la densità del fluido, e $\rho A(L)$ è la massa virtuale del fluido (per unità di lunghezza secondo x), che è influenzata dal moto, nel piano $x = L$, della sezione della pinna caudale col piano stesso, e che, con buona approssimazione, è

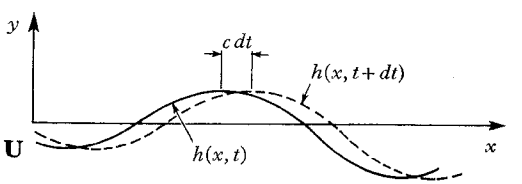


Figura 5.
Propulsione anguilliforme: configurazione dell'asse del pesce in due istanti successivi.

data dalla $\rho A(L) = \rho\pi b^2(L)/4$, se $b(L)$ è la dimensione della sezione trasversale sopra definita secondo l'asse y (normale al piano del moto ondulatorio). Infine la lineetta sopra un simbolo (a) indica il valor medio rispetto al tempo della grandezza rappresentata con detto simbolo, mentre l'indice scritto in basso nell'espressione tra parentesi significa che la grandezza tra parentesi deve essere calcolata per $x = L$.

L'espressione del rendimento del sistema propulsivo, definito come il rapporto fra la potenza necessaria, per produrre la spinta $\bar{T}$, e la potenza spesa $\bar{W}$ dall'organismo per compiere lo spostamento $h(x,t)$, è stata pure ottenuta da Lighthill: prendendo

$$(4) \qquad h(x,t) = H(x)F\left(t - \frac{x}{c}\right),$$

dove $F[t-(x/c)]$ è una combinazione lineare di $\cos\omega[t-(x/c)]$ e di $\sin\omega[t-(x/c)]$, di guisa che la (4) rappresenta un'onda che si propaga lungo l'asse con velocità c e lunghezza d'onda $\lambda = c/\omega$, detto rendimento risulta:

$$(5) \qquad \eta = 1 - \frac{1}{2}\frac{\left(1-\frac{U}{c}\right)^2 H^2(L)\bar{F}'^2 + U^2 H'^2(L)\bar{F}^2}{\left(1-\frac{U}{c}\right)H^2(L)\bar{F}'^2} \qquad F' = \frac{\partial F}{\partial t},\ H' = \frac{dH}{dx}.$$

Appare subito dalla (5) che è vantaggioso, per aumentare il rendimento, che sia

$$(6) \qquad H'(L) = \left(\frac{dH}{dx}\right)_{x=L} = 0$$

e pertanto che l'ampiezza dell'onda abbia un valore *estremo* in corrispondenza della pinna caudale: e questa proprietà è in accordo con quanto già segnalato, che l'ampiezza del moto ondulatorio osservata negli animali acquatici che praticano il sistema di propulsione in esame è *crescente* dalla testa alla coda. Con la condizione (6), l'espressione che dà la η si semplifica nella

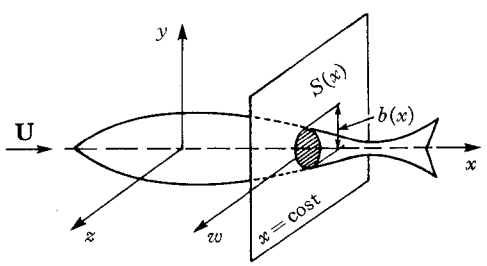


Figura 6.

Propulsione anguilliforme: applicazione della teoria dei corpi sottili alla determinazione della forza sul pesce.

$$(5') \qquad \eta = 1 - \frac{1}{2}\left(1 - \frac{U}{c}\right) = \frac{1}{2}\left(1 + \frac{U}{c}\right),$$

mentre la (3) dà

$$(3') \qquad \bar{T} = \frac{1}{2}\rho A(L)\cdot\left(1 - \frac{U^2}{c^2}\right)\bar{V}_L^2$$

essendo $V_L = (\partial h/\partial t)_L$ la velocità della parte estrema della pinna caudale nel moto ondulatorio. Appare da queste relazioni che affinché la spinta sia positiva, e pertanto il moto ondulatorio produca una propulsione e non una resistenza, deve essere $U < c$, e inoltre che nelle condizioni di massima efficienza la U deve avere il valore il piú possibile uguale a quello di c.

È interessante paragonare le relazioni ora indicate con quelle che dànno le grandezze analoghe nel caso del propulsore ad elica, ossia del sistema di propulsione in acqua quasi esclusivamente usato dai natanti fatti dall'uomo. Ora, indicando con c la velocità definita dalla $c = U + 2u$, in cui U è la velocità di traslazione dell'elica e $2u$ è l'incremento assiale di velocità all'infinito a valle indotto dall'elica stessa, dai teoremi della quantità di moto e dell'energia applicati considerando il moto entro il tubo di flusso costituito dalle linee di corrente, che si appoggiano ai punti del disco d'elica (ossia alla circonferenza di raggio uguale al raggio R dell'elica stessa), si ricava

$$(7) \qquad T = \frac{1}{2}\pi R^2\rho c^2\left(1 - \frac{U^2}{c^2}\right) \qquad \eta_a \simeq \frac{1}{2}\left(1 + \frac{U}{c}\right) \qquad \text{per } \frac{U}{c} \sim 1$$

che sono, almeno nelle condizioni di massimo rendimento, formalmente analoghe alle (3′) e (5′), cosí come si può riconoscere che la c, che appare nelle (7) ed è stata definita sopra, ha significato analogo a quello corrispondente alla (4).

Il rendimento del propulsore elicoidale dato dalla seconda delle (7) non rappresenta però il rendimento effettivo, ma soltanto quello che gli elicisti chiamano «rendimento assiale»: l'elica, ruotando, induce infatti anche una velocità angolare ω nel fluido, e l'energia cinetica corrispondente a questo moto è anch'essa energia perduta, in modo tale che, se si indica con η_p il cosiddetto «rendimento periferico» definito da $\eta_p = 1 - \omega/\Omega < 1$, il rendimento effettivo risulta $\eta = \eta_a\eta_p < \eta_a$.

D'altra parte, anche l'efficienza effettiva della propulsione anguilliforme può essere differente da quella data dalla (5′), per un fattore che è analogo a quello ora ricordato per il propulsore ad elica, a meno che la funzione $H(x)$ non soddisfi a determinate condizioni. In effetti, le condizioni di equilibrio dinamico – per le quali il risultante delle forze applicate ad ogni elemento del pesce in esame e delle forze d'inerzia di questo, e il momento risultante rispetto a un punto fisso generico delle stesse forze, debbono essere nulli – portano a scrivere due equazioni le quali (poiché le contrazioni muscolari grazie alle quali il pesce compie il moto ondulatorio descritto dalla (4) mettono in gioco solo forze interne, e quin-

di forze a risultante e a momento risultante nulli) contengono solo le funzioni $S(x)$, definente la legge di variazione dell'area delle sezioni trasversali del pesce, e $h(x,t)$, definente il moto ondulatorio ed espressa dalle (4). Solo se la $H(x)$ è tale che le predette equazioni siano soddisfatte, l'efficienza del sistema ha proprio il valore (5′); in caso contrario, mentre le contrazioni muscolari producono il moto ondulatorio, che dà luogo alla traslazione secondo x, a questo moto si sovrappone anche un moto rigido di deriva secondo l'asse z, e un moto rotatorio intorno all'asse y, di guisa che la traiettoria del centro di gravità non risulta piú rettilinea, ma assume un andamento sinuoso. In queste condizioni l'efficienza è ovviamente minore di quella data dalla (5′), come conseguenza dell'energia corrispondente a questi moti parassiti, che è dissipata, o almeno non utilizzata ai fini della traslazione secondo x. Se ora si ricerca a quali condizioni deve soddisfare la funzione H affinché questi moti parassiti *non* si producano, si ottiene che esse s'identificano con le condizioni esprimenti che nel moto ondulatorio corrispondente il sistema è dinamicamente equilibrato, ossia che il risultante e il momento risultante delle forze d'inerzia del corpo prodotti da detto moto sono *nulli*. Esplicitando la funzione H, cosí da poter scrivere le equazioni di equilibrio dinamico, si ricava la seguente importante proprietà: determinata la forma della funzione d'onda H, per la quale le sopraddette equazioni sono soddisfatte per una data velocità U di avanzamento, se la lunghezza d'onda del moto ondulatorio in questione rimane costante, allora, al variare di U e per tutti i valori di detta velocità per i quali il coefficiente di resistenza C_r (coefficiente definito dalla $F_r = C_r \rho S U^2/2$, in cui F_r è la resistenza del corpo traslantesi con la velocità U, mentre S è la superficie di riferimento del corpo stesso), può essere considerato costante, la sua configurazione non cambia e il rendimento della propulsione rimane costante; è solo la frequenza ω del moto ondulatorio che cambia, e cambia proporzionalmente a U. La regolazione a lunghezza d'onda costante appare pertanto particolarmente semplice e vantaggiosa, e d'altra parte essa è anche la piú conveniente nei riguardi della struttura scheletrica dell'organismo, mentre le variazioni del coefficiente di resistenza C_r ai differenti valori di U possono essere facilmente controbilanciate da una corrispondente variazione soltanto dell'ampiezza h_0 dell'oscillazione della pinna caudale (e precisamente, facendo variare h_0 proporzionalmente a $C_r^{1/2}$). È appena il caso di porre in evidenza la grande elasticità di adattamento alle varie condizioni del moto, e la notevole finezza che il sistema di regolazione ora indicato ha in paragone agli analoghi artifici inventati dall'uomo per cercare di ottenere risultati analoghi, ma senza ottenerli in modo cosí perfetto, per il sistema di propulsione ad elica (ad esempio elica a passo variabile in moto).

2.5. Propulsione carangiforme.

Appare dalla relazione (3′) che la spinta dipende solo dalle condizioni del moto ondulatorio all'estremo della pinna caudale; è pertanto logico domandarsi se è proprio necessario che l'intero corpo del pesce prenda parte al moto ondulatorio. La risposta a tale quesito deriva subito dalle considerazioni svolte nel

paragrafo precedente, in quanto le condizioni che debbono essere soddisfatte affinché il sistema sia dinamicamente equilibrato mettono in evidenza che la forma dell'onda, richiesta per detto equilibrio, dipende dalla forma del corpo. Ora, per i pesci che usano la propulsione anguilliforme la forma è tale da rendere necessario che l'ondulazione sia estesa a tutta la lunghezza, anche se l'ampiezza aumenta dalla testa alla coda. Però, cambiamenti morfologici possono permettere di ridurre l'ondulazione a una porzione ristretta del corpo vicino alla pinna caudale, intendendo con questo dire che l'ampiezza dell'onda è dappertutto molto piccola in confronto a quella della parte sopra indicata. Si ha, in questo modo, l'evoluzione dalla propulsione anguilliforme a quella carangiforme, cosiddetta perché è quella usata dai pesci della famiglia Carangidi (come il fanfano, la seriola, la leccia), ma anche impiegata dai pesci di diverse altre famiglie, come quella delle Clupeidi (ad esempio l'aringa), dei Seriamidi (ad esempio il branzino), ecc.

Questi pesci hanno la capacità di compiere *performances* alquanto piú elevate di quelle dei pesci che usano il modo anguilliforme; in particolare raggiungono velocità massime, che, misurate in (lunghezza del corpo per secondo) variano dai 6ℓ/s ai 12ℓ/s. La caratteristica essenziale della loro forma è la seguente: si chiami, per analogia con la nomenclatura usata per i profili alari, «spessore», la dimensione del profilo del pesce (inteso come la sezione del corpo del pesce nel piano di simmetria di questo) misurata normalmente alla corda del profilo stesso. Ora, questo spessore ha un valore minimo, che è relativamente molto piccolo, nella parte del corpo immediatamente prima della pinna caudale, e proprio in corrispondenza di questa l'ampiezza del moto ondulatorio incrementa molto rapidamente, mentre a monte di essa vi è una regione a spessore relativamente grande, e ad ampiezza di ondulazione relativamente molto piccola (fig. 7).

La relazione che dà la spinta di propulsione nel sistema carangiforme differisce da quella data per il sistema anguilliforme per alcune variazioni apportate da Wu [1971] per tener conto dell'influenza dello strato vorticoso nella scia delle pinne ventrale e dorsale, e da Lighthill [1971] per tener conto del fatto che l'ampiezza delle oscillazioni della pinna caudale può, nel sistema carangiforme, assumere valori alquanto piú grandi di quella corrispondente al sistema anguilliforme. Non è il caso di riportare qui dette variazioni; quello che importa invece osservare è che il confronto tra i risultati sperimentali (ottenuti da Bainbridge e re-

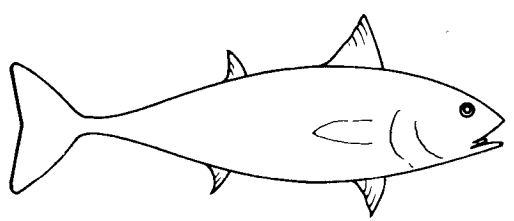


Figura 7.
Forma tipica dei pesci con propulsione carangiforme.

lativi al Leucisco della famiglia dei Ciprinidi), e i risultati calcolati da Lighthill darebbe un valore del coefficiente di resistenza C_r quattro volte piú grande di quello corrispondente, per il medesimo numero di Reynolds, a un solido di rivoluzione di pari forma, quale risulta dalle misure in galleria idrodinamica.

Un risultato analogo è stato pure ottenuto da Smith, che, stimando la potenza spesa dai muscoli del *Carassius auratus* della stessa famiglia dei Ciprinidi dal consumo di energia durante il moto, e supponendo un rendimento propulsivo di 0,8, dedusse che la potenza spesa da detto pesce era 3,6 volte piú grande di quella calcolata per il caso del moto rigido di un corpo di forma corrispondente e per lo stesso numero di Reynolds.

Per spiegare questo risultato fu avanzata da Lighthill [1971] e altri la congettura che l'aumento di resistenza sia dovuto a un incremento della resistenza di attrito viscoso connessa con una diminuzione dello spessore dello strato limite adiacente alla superficie lambita, provocata dal moto ondulatorio. Questa congettura può essere giustificata assimilando l'influenza di detto moto ondulatorio sulla resistenza a parete con quella corrispondente che è prodotta dall'aspirazione e dal soffiamento dello strato limite, calcolata ad esempio da Pretsch ed altri. Si devono d'altra parte, a questo riguardo, fare due osservazioni: l'una è che l'incremento di resistenza accennato è relativamente di tanto minore importanza quanto minore è il rapporto $U/(\overline{V}_L^2)^{1/2}$, e poiché U e $(\overline{V}_L^2)^{1/2}$ crescono entrambi proporzionalmente a ω, mentre U è proporzionale a L e $(\overline{V}_L^2)^{1/2}$ all'ampiezza h_0 dell'oscillazione, si comprende la convenienza per il pesce di rendere tale ampiezza la piú grande possibile. Inoltre, per diminuire l'effetto dell'incremento di resistenza sopra indicato, può essere conveniente per il pesce alternare periodi in cui sviluppa una spinta relativamente elevata, che gli imprime una velocità maggiore di quella desiderata, con periodi senza propulsione, durante i quali diminuisce la velocità ma è nullo l'incremento di resistenza d'attrito dovuto al moto ondulatorio.

La seconda osservazione è che la medesima causa che può produrre un incremento della resistenza viscosa, e cioè il minor spessore dello strato limite per un dato numero di Reynolds, può d'altra parte influire sul fenomeno della transizione da regime laminare a regime turbolento, aumentando la stabilità del flusso laminare, e pertanto determinando un incremento del numero di Reynolds per il quale la transizione stessa si produce. Questo può fare intuire che per i pesci piú veloci, per i quali il numero di Reynolds per la massima velocità può essere alquanto piú grande di 10^6, il fenomeno ora discusso possa non presentarsi, o almeno produrre effetti trascurabili.

2.6. Propulsione a pinna battente.

I valori della spinta relativamente piccoli, che anche col sistema carangiforme è possibile ottenere, non permettono di raggiungere velocità molto elevate; in effetti si riconosce che, col crescere della capacità di nuotare a grande velocità, si verifica una graduale modificazione della pinna caudale, che per i pesci piú veloci ($30 \div 40$ km/h per un tempo relativamente lungo) prende la configurazione di un'ala rastremata di grande allungamento (a forma di lunula, da cui la denominazione «carangiform propulsion with lunate tail» usata da Lighthill). Questo modo di propulsione è usato da parecchie famiglie dell'ordine Percomorfi (come la famiglia degli Scombridi, cui appartiene lo scombro; quella dei Tunidi, come il tonno; quella degli Xifidi, come il pesce spada) ed è analogo a quello degli uccelli, con la differenza che le ali degli uccelli non solo debbono provvedere a dare la necessaria spinta per avanzare, ma anche la *portanza* necessaria per sostenere il peso, mentre le pinne dei pesci, almeno quelli appartenenti alle famiglie indicate, debbono solo propellere.

Lo studio di questo sistema di propulsione non può essere fatto con la teoria dei corpi sottili, come è stato fatto per i due precedenti modi presi in considerazione; si deve, appunto come per gli uccelli, applicare la teoria delle superfici portanti in moto non stazionario. Parecchie ricerche al riguardo sono state fatte particolarmente da Lighthill, da Wu, da Newman e Chopra, ma di esse ci si limita a darne qui cenno, perché troppo specializzate in argomento, e, a parte la complicazione, ancora troppo incomplete per portare a risultati che diano un certo affidamento. Possono invece essere di qualche interesse le seguenti semplici considerazioni.

La spinta è una conseguenza della variazione di quantità di moto impressa al fluido nella direzione dell'asse x, e prodotta dai vortici nella scia generata dall'azione della pinna nel suo moto oscillatorio: piú precisamente, sono i vortici, il cui asse è nella direzione dell'asse y (fig. 8) ad indurre tale variazione. Ora gli assi di detti vortici possono essere considerati disposti come le generatrici della superficie cilindrica, la cui direttrice nel piano (x,z) è la sinusoide, che è descritta dal bordo di fuga della pinna alare posto nello stesso piano e la cui ampiezza è uguale a h_0. La massa del fluido che è influenzata da detti vortici può essere considerata proporzionale alla quantità $\rho U b h_0$, se b è l'apertura della pinna (ossia, la sua massima dimensione nella direzione y), di guisa che se u_1 rappresenta un opportuno valore medio della velocità indotta dai vortici sopra indicati all'infinito a valle, nella direzione dell'asse x, la variazione della quantità di moto di detta massa, ossia la spinta, può essere espressa con la

$$\overline{T} = \rho \chi U b h_0 u_1 \tag{8}$$

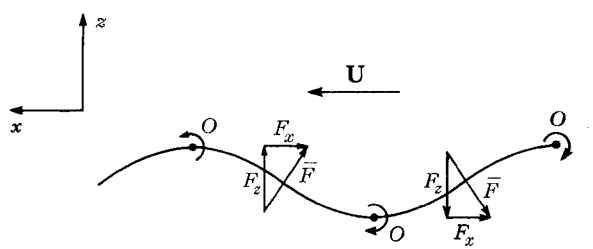


Figura 8.
Vortici generatori della spinta nella propulsione a pinna battente.

essendo χ un fattore di proporzionalità. D'altra parte, procedendo in modo analogo a quanto si fa per il propulsore elicoidale, si ricava che il rendimento del sistema è ancora dato da una formula analoga alla (7), e precisamente

$$(9) \qquad \eta = \frac{1}{1+\frac{u}{U}}$$

se $u = u_1/2$. Esprimendo allora $\bar{T}$ con un coefficiente di carico C_T dato dalla

$$(10) \qquad C_T = \frac{\bar{T}}{\frac{1}{2}\rho U^2 S_p}$$

in cui S_p è la superficie della pinna alare, si ricava

$$(9') \qquad \eta = \frac{1}{1+C_T\left(\frac{l_p}{h_0}\right)\left(\frac{1}{4\chi}\right)}$$

se l_p è la lunghezza della corda di detta pinna corrispondente alla $S_p = bl_p$. Appare pertanto che il rendimento propulsivo è tanto maggiore quanto minore è il coefficiente di carico C_T, e, a parità di C_T, quanto maggiore è l'ampiezza dell'oscillazione rispetto alla corda della pinna alare.

Il raggiungimento di forti valori dell'efficienza propulsiva, prossimi a 1, appaiono possibili per il sistema in esame; ma, se questa proprietà è una condizione necessaria per il raggiungimento delle alte velocità, come quelle sopra indicate, essa non è ancora una condizione sufficiente. Si deve ancora riconoscere se la potenza richiesta a tale velocità può essere effettivamente sviluppata dal pesce.

Ora, ad esempio Focke, in una nota sulla «Zeitschrift für Flugwissenschaft» del 1965 (vol. IV), riporta per un delfino i seguenti dati: lunghezza m 2,75; rapporto di finezza 1/6; peso 90 kg; velocità 10 m/s; superficie bagnata 1,98 m^2; numero di Reynolds $2{,}7 \times 10^7$; coefficiente di resistenza $C_r = 0{,}003$. Ne risulta una resistenza $F_r = 30{,}3$ kg e una potenza corrispondente di 303 kgm/s; prendendo un rendimento propulsivo, $\eta = 0{,}9$, la potenza meccanica che deve essere sviluppata risulta 335 kgm/s = 4,5 HP. Stimando che per un peso di 90 kg il corrispondente peso dei muscoli che producono il moto sia di 36 kg, si ottiene una potenza per chilogrammo di muscolo 0,12 HP/kg = 0,054 HP/lb.

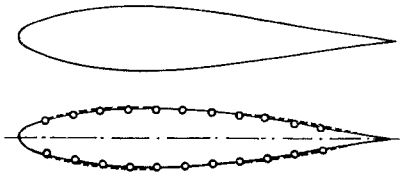


Figura 9.
Paragone del profilo della trota e del profilo NACA 63/4016.

Focke, sulla base di informazioni date dal biologo Friedrich, direttore dello Übersee di Brema, stima che un delfino possa sviluppare una potenza specifica di 5 kgm/s = 0,03 HP/lb, valore nettamente inferiore a quello sopra indicato. Si deve perciò ammettere che i pesci, almeno quelli di certe famiglie, hanno un coefficiente di attrito che, a parità di numero di Reynolds, è alquanto minore di quello dei corpi rigidi simili. A questo riguardo è da ricordare che nel suo mirabile volumetto *Aerodynamics* Kármán ricorda le ricerche pionieristiche di George Cayley sulla forma idrodinamica ottima (nei riguardi della resistenza) e cita quanto egli scrisse in proposito nel suo *Aeronautical Miscellaneous Notebook*: «Io penso che questo problema è cosí difficile che è molto piú utile investigarlo per mezzo di esperimenti che per mezzo di speculazioni teoriche; e poiché non ho nessun risultato conclusivo a mia disposizione a questo riguardo, la sola cosa da fare è copiare la natura: perciò io ho preso come forma ottima quelle slanciate della trota e della beccaccia» [1956, trad. it. p. 54].

Kármán riproduce pure il profilo della trota ottenuto e disegnato da Cayley, e paragona questo profilo a quelli dei profili alari «laminari» (cosí detti perché essi, per la distribuzione delle pressioni sul contorno cui dànno luogo, producono uno spostamento a valori piú elevati del numero di Reynolds critico per la transizione al regime turbolento) NACA 63/4016 e LBN-0016. Ora, è sorprendente la coincidenza del profilo della trota con quello dei profili sopra citati (fig. 9). Un analogo confronto è stato fatto da Hertel [1963] nel suo trattato *Struktur-Form-Bewegung* per il delfino e per il profilo alare NACA 66/018, e ancora una volta la coincidenza dei profili appare sorprendente, e, ciò che è interessante osservare, le piccole differenze che si notano sono tutte a vantaggio del delfino (fig. 10), particolarmente nella profilatura della parte poppiera.

Questo effetto di forma non è però sufficiente a spiegare il fenomeno sopra indicato, e che spesso è indicato come «paradosso di Gray». Sono state fatte perciò da diversi autori varie ipotesi: Kramer in un articolo apparso nel «Journal of the American Society of Naval Engineers» (1960) attribuí la riduzione dell'attrito a un effetto viscoelastico della parete, e alla particolare struttura della pelle del delfino, che, osservata al microscopio, appare essere costituita da uno

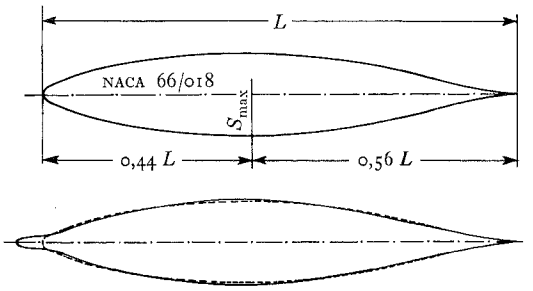


Figura 10.
Paragone del profilo del delfino e del profilo NACA 66/018.

strato interno relativamente rigido, che fa da supporto a strati cedevoli, nei quali sono disposti parecchi condotti contenenti materiali spugnosi. Poiché la quantità di acqua in questa sostanza spugnosa è notevole, Kramer avanzò l'ipotesi che gli strati esterni si comportino come una parete flessibile, che trasferisce al fluido nei condotti interni le oscillazioni di pressione del fluido esterno dovute a instabilità del flusso laminare; il liquido che fluisce in questo dissipa l'energia corrispondente a dette oscillazioni, e come effetto di questo assorbimento di energia ne deriverebbe una stabilizzazione del flusso laminare.

I risultati degli esperimenti per controllare la validità dell'ipotesi di Kramer non sono fra loro concordanti, se non in questo, che se effettivamente le proprietà viscoelastiche della parete possono produrre una riduzione dell'attrito, questa non è certamente dovuta a una stabilizzazione del regime laminare.

Secondo un'altra ipotesi, avanzata già da Breder nel IV volume di «Zoologica» (1926), la riduzione dell'attrito sarebbe prodotta dal muco che è secreto dalla pelle del pesce, ma questa congettura non fu allora appoggiata da ricerche sperimentali o teoriche, fino a che recentemente Rosen e Cornfold [1971] resero noti alcuni risultati molto interessanti. Questi autori presero alcuni campioni *vivi* di parecchie famiglie di pesci; immediatamente dopo averli tolti dall'acqua, fecero uscire dai loro corpi l'eccesso di acqua di cui erano imbevuti ed estrassero campioni dell'umore viscoso secreto dalla loro pelle. Dopo avere di-

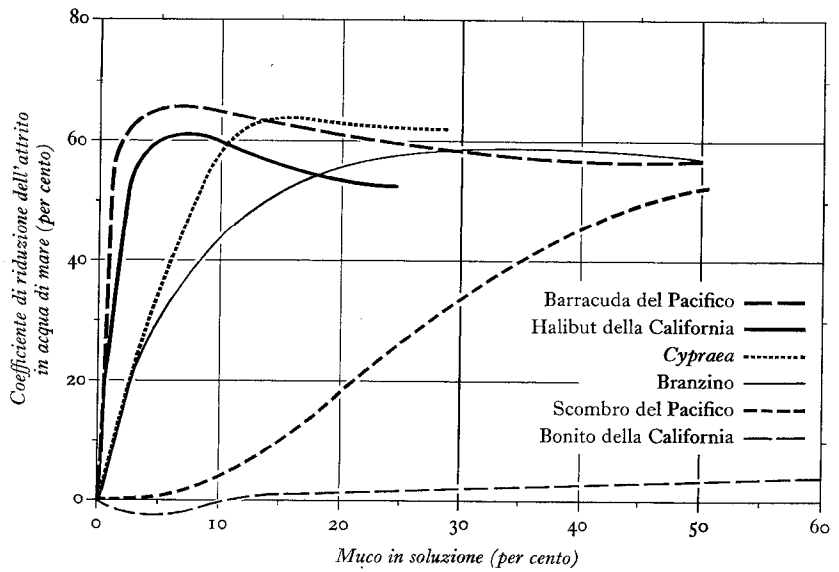


Figura 11.

Riduzione dell'attrito, in acqua salata, prodotto dal muco secreto dai pesci (secondo Rosen e Cornfold).

luito questo umore, cosí da avere soluzioni di questo sia in acqua dolce sia in acqua marina con diversi valori della concentrazione, compirono con le varie soluzioni una serie di esperimenti col reometro di Hoyt. I risultati di questi esperimenti sono mostrati nelle figure 11 e 12, nelle quali sono riportate le riduzioni percentuali dei coefficienti di attrito in funzione della concentrazione delle soluzioni per diversi tipi di pesci. Poiché il numero di Reynolds degli esperimenti era circa 8000, non vi è dubbio che il regime di flusso nel reometro fosse turbolento. Ora, si può riconoscere che per quasi tutti i campioni sperimentati la riduzione dell'attrito è notevole; essa è massima per il barracuda del Pacifico (*Shyraena argentea*), per il quale detta riduzione raggiunge il valore percentuale del 65,9 per cento, per una concentrazione della soluzione del 5 per cento: questi risultati riguardano prove fatte con campioni della stessa specie, catturati in diversi giorni e aventi lunghezza da 70 a 80 cm. Pressoché uguale efficienza come attrito-riduttore presentano l'umore dello halibut e, tra i pesci d'acqua dolce, del pesce persico e della trota. La sola importante eccezione è mostrata dal bonito della California (*Sarda chilensis*), che pure è uno dei piú veloci nuotatori, mentre la *Cypraea spadicea*, che non è un pesce, ma un mollusco, e non ha bisogno di nuotare, secerne un umore molto efficiente.

Gli sperimentatori sopra citati hanno messo in evidenza due proprietà: prima, l'umore secreto dalla pelle non è solubile in acqua tranquilla, mentre di-

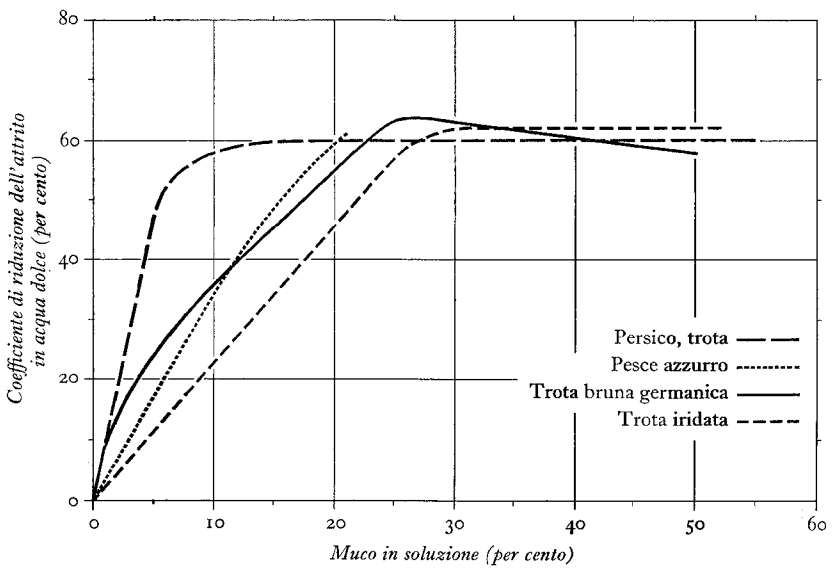


Figura 12.

Riduzione dell'attrito, in acqua dolce, prodotto dal muco secreto dai pesci (secondo Rosen e Cornfold).

venta solubile se l'acqua è violentemente agitata; in secondo luogo, la quantità di materiale organico che il pesce deve perdere per ridurre fortemente l'attrito è piccolissima: cosí ad esempio il barracuda per ridurre del 44 per cento la resistenza deve introdurre nel fluido dello strato limite non piú di 5,7 parti per milione delle sue sostanze organiche.

Rosen e Cornfold proposero di conseguenza le seguenti ipotesi: per la maggior parte del tempo il pesce nuota con velocità abbastanza piccola perché il flusso nello strato limite sia laminare, e in queste condizioni, essendo il muco non solubile, lo straterello di esso adiacente alla pelle rimane attaccato al corpo. Ma nei brevi intervalli di tempo, nei quali il pesce deve nuotare rapidamente per afferrare una preda, o per sfuggire da un predatore, lo strato limite diventa turbolento; una porzione, molto piccola, come si è detto, del muco può sciogliersi in acqua in conseguenza del forte rimescolamento turbolento, e cosí le condizioni per ottenere una drastica riduzione dell'attrito sono soddisfatte, e il pesce può raggiungere velocità elevate con una relativamente modesta spesa d'energia, e con una piccolissima perdita di sostanze organiche.

Benché questa ipotesi sia basata su dati sperimentali accurati, essa non è ancora accettata da tutti, il che non stupisce; comunque, benché sia certo che ancora parecchie ricerche debbano essere eseguite da biologi, chimici, e idromeccanici prima di arrivare a una spiegazione completa del fenomeno, sembra indubbio che i risultati sopra riportati siano di grande utilità per indicare la via da seguire per arrivare a detta spiegazione.

2.7. Propulsione a getto.

Se il sistema di propulsione piú largamente usato dagli animali acquatici è quello a moto ondulatorio, sia pure realizzato nei differenti modi che sono stati specificati nei paragrafi precedenti, vi sono peraltro parecchie altre forme di propulsione proprie di determinate specie. Cosí ad esempio per la classe dei Cefalopodi, e in questa per i calamari (*Loligo*), per gli Ottopodi come i polpi, per le seppie, la spinta è ottenuta eiettando getti d'acqua in rapidi ripetuti impulsi: essi realizzano, in altre parole, la propulsione a getto pulsante. Lane dice a questo riguardo in *Kingdom of the Octopus* (1960): «Ere prima che l'uomo scoprisse la propulsione a getto, Cefalopodi come i calamari si propellevano a getto nei mari primordiali. Alcune delle piú piccole specie di calamari mostrano i migliori esempi di questi propulsori a getto naturali mentre per i loro rapidi movimenti si meritarono nomi come quelli di "freccia del mare" o di "calamari volanti" (*Onychotenthis*, ecc.)» [citato in Siekmann 1963, p. 399].

In effetti è stato osservato che un calamaro come la *Loligo*, della massa di 100 g, con un singolo impulso può accelerare dallo stato di riposo a una velocità di 2 m/s, mentre per altri calamari si vuole da alcuni che possano raggiungere velocità massime fino a 8 ÷ 9 m/s (si noti che la velocità massima raggiunta dall'uomo nuotando su un percorso di 100 m è di circa 2 m/s, e correndo su pista è di circa 10 m/s).

Il sistema usato per la propulsione è simile a quello degli aerei a getto: l'acqua viene aspirata (fig. 13) attraverso un'apertura (presa d'acqua), disposta sul mantello in corrispondenza del collo (ossia nella regione di giunzione della testa al resto del corpo), per effetto di un'espansione della parete del mantello stesso, e penetra in una cavità di questo, che agisce come camera di compressione: durante questa fase, il condotto di emissione è parzialmente collassato e chiuso. Successivamente la presa d'acqua viene chiusa ermeticamente grazie a un moto che fa ritrarre la testa verso il corpo; questo porta piccole costole nell'interno del mantello, che in conseguenza del movimento sopraddetto penetrano entro opportune sedi cave poste sul lato opposto, funzionando cosí come la valvola di chiusura dei motori a combustione interna. La parete della cavità del mantello, dotata di una robusta muscolatura, si contrae quindi violentemente e spinge l'acqua in essa contenuta ad uscire con grande velocità attraverso il condotto di emissione, il cui asse può essere orientato dall'animale a piacimento. L'intero ciclo ha la durata di circa un secondo.

Quanto sopra è stato detto vale per le specie piú veloci e nella figura 13 è indicato lo schema del sistema ora accennato per un comune calamaro; per quelle piú lente, come gli Ottopodi, il sistema si semplifica, cosí che in alcuni Ottopodi (*Chuniotenthis*) si ha un solo condotto che funziona sia per l'aspirazione sia per lo scarico, mentre nei nautili, che sono tra i Cefalopodi i piú primitivi, manca anche la camera di compressione e l'acqua viene aspirata e scaricata in conseguenza di espansioni e contrazioni del solo condotto sopra indicato.

L'origine della spinta è sempre, come già si è detto, nella capacità dell'animale di produrre una variazione della quantità di moto del mezzo in cui vive: un confronto tra il sistema ora considerato e quelli indicati nei paragrafi precedenti in base al criterio dell'efficienza non può farsi perché mancano troppi elementi sperimentali e teorici al riguardo: in particolare, il getto scaricato dal condotto di emissione è generalmente assai prossimo al corpo dell'animale ed interferisce sul flusso attorno a questo nel senso di diminuirne la resistenza, e un effetto analogo è prodotto dalla contrazione della parete della cavità del mantello,

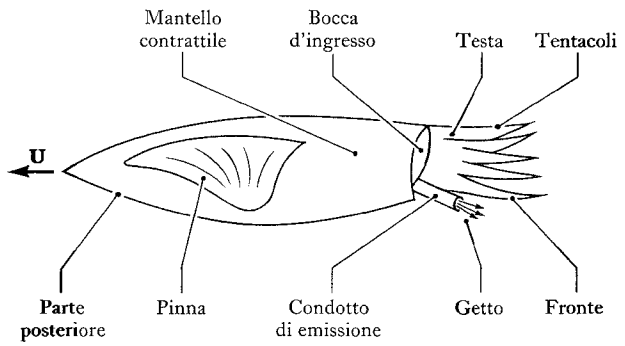


Figura 13.
Schema del sistema di propulsione a getto della *Loligo*.

in conseguenza della diminuzione dell'area della sezione trasversale del corpo che ne deriva. D'altra parte un confronto basato sul solo criterio dell'efficienza probabilmente non ha un gran senso, perché altri elementi, morfologici, fisiologici, condizioni di vita, ecc. possono determinare la preferenza di un sistema rispetto a un altro.

2.8. Propulsione a pale e a remi.

Questo sistema di propulsione, in forma molto perfezionata rispetto a quella, in paragone, rudimentale ideata dall'uomo è praticato da molti animali acquatici, in particolare i crostacei. Tra questi i granchi, alla cui specie appartengono alcuni dei migliori nuotatori, quali i Portunidi, che hanno la particolarità di potersi muovere sia in direzione trasversale, rispetto al corpo, sia in direzione longitudinale, grazie all'azione dell'ultimo paio di zampe, che all'estremo libero terminano a forma di pale ovali, appiattite e qualche volta leggermente incavate a guisa di cucchiaio. Nel moto trasversale (che è il piú efficiente) le due pale battono l'acqua avanti e indietro ritmicamente e sincronicamente in piani ad angolo retto con la direzione del moto, facendo descrivere al baricentro di ogni pala una traiettoria quasi circolare, e ruotando le pale stesse leggermente, nel passaggio dalla corsa in avanti a quella di ritorno, in modo che entrambe risultino efficienti al fine della propulsione (fig. 14): appare pertanto che le zampe funzionano come eliche a pale orientabili montate sopra un corpo fusiforme, che presenta uguale profilatura per gli spostamenti laterali in entrambi i sensi. Per i granchi della specie *Polybius*, aventi una larghezza di 15 cm e pale di corda della lunghezza di 1,5 cm, battenti con una frequenza di 10,5 Hz, sono state misurate velocità di 1,33 m/s, ma è certo che essi possono nuotare con velocità molto piú elevate, come è dimostrato dal fatto che essi possono raggiungere e catturare scombri in mare aperto.

Del tutto diverso è il sistema usato nel caso di moto in direzione longitudinale: le pale, destra e sinistra, in questo caso sono mosse non sincronicamente, ma alternativamente, ed agiscono come le pale di un remo, che sono attive soltanto durante la corsa di lavoro, che è quella che avviene dal dietro all'avanti (co-

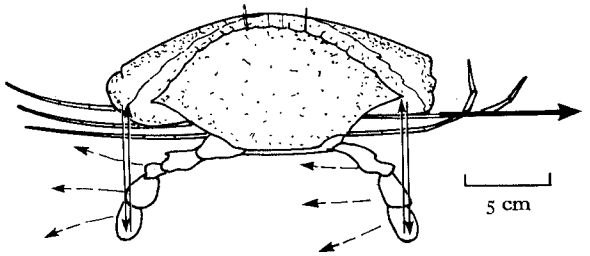


Figura 14.
Granchio portunide con propulsione a pale.

sicché il senso del moto è sempre retrogrado); probabilmente anche in questo caso si ha un effetto di interferenza favorevole tra la corrente d'acqua prodotta dall'azione del remo e il corpo dell'animale.

Altri crostacei, come le aragoste, utilizzano per la propulsione rapidi, ripetuti colpi di coda, ottenuti con inflessioni ventrali di questa, e che agiscono come il battere di un remo: la direzione del moto è ovviamente sempre longitudinale, e il senso del moto quello retrogrado. L'efficienza deve essere notevole, nonostante l'apparente grossolanità dell'azione, dal momento che velocità fino a 1,2 m/s sono state frequentemente misurate, e si crede che, almeno per la specie *Palinurus*, la velocità massima possa arrivare fino a 8 m/s.

Al lato opposto, per quanto si riferisce alle dimensioni, è il caso della specie *Diaptomus* della lunghezza di circa 2 mm, che può fare brevi scatti di velocità fino a valori di questa di 20 cm/s, ossia fino a cento volte la lunghezza del suo corpo per secondo, prestazione questa veramente eccezionale. Il suo apparato motore è costituito da cinque paia di zampe, poste sul torace, e aderenti a questo quando l'animale galleggia in riposo sull'acqua. Le prime quattro paia di zampe sono frangiate di specie di setti piumacei con setole uniformemente spaziate. Queste costituiscono come delle pale a fessure, che mentre sono molto efficienti per produrre un incremento dell'azione idrodinamica presentano piccola inerzia, potendosi inoltre spiegare a ventaglio durante la corsa di lavoro, e ripiegarsi quindi aderenti alla costola portante durante quella di ritorno (fig. 15).

Per produrre la propulsione, le zampe sono spinte violentemente dal torace verso la coda, agendo come remi, e ciascun paio iniziando la sua corsa dopo che il paio precedente l'ha compiuta per intero; nello stesso tempo, le antennule, portate in corrispondenza della testa, che durante la fase di galleggiamento a riposo si estendevano lateralmente turgide e rigide, e che pochi millisecondi prima del balzo delle zampe remiganti erano diventate floscie, si ripiegano passivamente lungo il corpo. Quando tutte le cinque coppie di remi hanno compiuto la loro corsa di lavoro, esse ritornano insieme nella posizione di partenza ripiegando sul corpo in modo da rendere minima la loro azione sull'acqua. La durata dell'intero ciclo è stata stimata da Storch [1929] in 1/72 di secondo: nell'ipotesi che la corsa di lavoro e quella di ritorno abbiano la stessa durata, ne deriverebbe, per

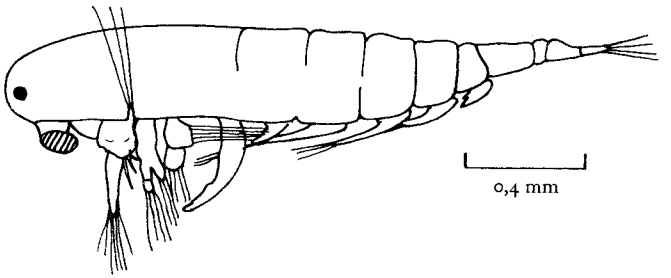


Figura 15.
Propulsione a remi del *Diaptomus* per l'avanzamento a scatti.

una velocità di 20 cm/s, un'accelerazione di circa tre volte quella di gravità!

Oltre che prodursi in spunti di velocità cosí elevata, come quella ora detta, il *Diaptomus* può pure nuotare con velocità costante e notevolmente piú bassa, usando un sistema propulsivo del tutto diverso da quello sopra descritto: apparentemente a conoscenza della teoria dei vortici, esso produce, con un movimento rapido rotatorio delle antennule prima menzionate, due vortici di circuitazione uguale e contraria, ai due lati del corpo, i quali inducono nell'acqua una variazione di quantità di moto in direzione longitudinale, e quindi, per reazione, una spinta sull'animale (fig. 16).

Da notare che mediante altre appendici portate dal corpo di questo, e di molto minore estensione delle antennule, viene prodotto un altro paio di vortici uguali e contrari nell'immediata prossimità del corpo stesso, e di minore intensità di quelli prima indicati: la corrente da essi indotta ha essenzialmente lo scopo di portare il cibo alle mandibole.

Un sistema di propulsione analogo a quello ora indicato per il *Diaptomus* è usato dal gamberetto di mare della specie *Chirocephalus*: questo, che ha una lunghezza di circa 2,5 cm, generalmente nuota sul dorso, e può sia muoversi avanti e indietro, sia librarsi in un dato posto (pure avendo densità superiore a quella dell'acqua), sia innalzarsi o abbassarsi (fig. 17). A tale scopo esso è provvisto di undici paia di zampe, che può muovere avanti e indietro con ritmo metacronale, come remi di una barca, sfruttando inoltre l'azione di pompaggio derivante dalle variazioni di volume degli spazi compresi tra zampa e zampa, e tra queste e il corpo, mentre alcune paia di dette zampe, provviste di setti piumosi (exopoditi) contemporaneamente possono compiere movimenti rotatori, atti a generare vortici dello stesso tipo di quelli indicati per il *Diaptomus*; è stata osservata una frequenza di battito delle zampe remiganti fino a 12,5 Hz, e una velocità di traslazione fino a 14 cm/s (lunghezza dell'animale circa 2,5 cm).

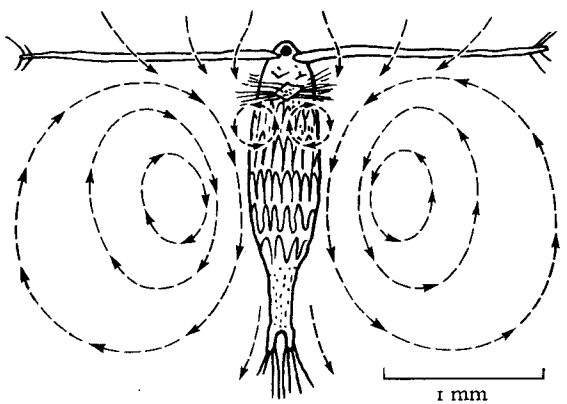


Figura 16.
Sistema vorticoso nell'avanzamento continuo del *Diaptomus*.

Figura 17.
Rappresentazione schematica delle zampe di propulsione del *Chirocephalus*.

2.9. Altre proprietà del moto dei pesci.

È già stata messa in evidenza, in qualche caso, la capacità che hanno diversi animali acquatici di imprimere al proprio corpo accelerazioni elevatissime: in effetti, questa proprietà è posseduta da diverse specie di pesci, ed in particolare da quelli che hanno pinne caudali molto sviluppate. Cosí, in particolare, è stata misurata un'accelerazione, nella prima fase di moto a partire dallo stato di quiete, fino a 50 m/s^2 (pari quindi a circa 5 g) per i lucci, e per le trote fino a 40 m/s^2 (ossia circa 4 g).

Dall'esame fotografico della fase di partenza dalla quiete [Weish 1973] è stato visto che tutta la manovra è fatta in tre fasi: nella prima fase il pesce, che inizialmente è col suo asse rettilineo, si dispone in modo che la coda sia quasi ad angolo retto col resto del corpo, formando cosí col suo asse una linea a L; nella seconda fase, la coda ruota agendo sull'acqua con forza che è proporzionale all'accelerazione angolare di detto moto, e ricavandone quindi per reazione la forza propellente; nella terza fase il pesce raddrizza il proprio asse ruotando in una direzione che forma con la direzione originaria (ossia all'istante iniziale) un angolo proporzionale all'accelerazione impressa sopraddetta (corrispondente alla rotazione di insieme del pesce prodotta dal momento della forza di reazione idrodinamica). Del tutto analoga alla proprietà ora indicata è la capacità, che parecchie specie di pesci hanno in modo davvero stupefacente, di potere variare la direzione del loro moto: in effetti si possono avere traiettorie con raggi di curvatura uguale, o anche minore, alla lunghezza del pesce e con apparente minima perdita di energia. Anche in questo caso lo sviluppo della pinna caudale appare essere di primaria importanza, ed anche ora l'intero processo appare risultare da tre diverse fasi caratterizzate dal diverso tipo di moto del centro di gravità dell'animale. Nella prima fase detto centro continua a muoversi di moto rettilineo nella direzione originaria; nella seconda esso descrive un arco, che è molto approssimativamente un arco di cerchio, mentre nella terza fase finale ritorna a muoversi di moto rettilineo nella nuova direzione.

La molteplicità dei sistemi di propulsione e la perfezione con la quale ciascuno di detti sistemi è stato realizzato mostrano che se l'uomo non ha potuto dare ragione del *perché* dei fenomeni di moto, la natura ha ben saputo *come* produrre i fenomeni stessi; a conforto dell'uomo, si può dire che essa ha avuto piú tempo a sua disposizione. [C. F.].

Accademia del Cimento
1667 *Saggi di naturali esperienze fatte nell'Accademia del Cimento*, Cocchini, Firenze.

Berg, M. C., e Anderson, R.
1973 *Bacteria swim by rotating their flagellar filaments*, in «Nature», CCXLV, pp. 380-82.

Berg, M. C., e Turner, L.
1979 *Movement of micro-organisms in viscous environments*, in «Nature», CCLXXVIII, pp. 349-51.

Blake, J.
1972 *A model for the micro-structure in ciliated organisms*, in «Journal of Fluid Mechanics», LV, parte I, pp. 1-23.

Born, M.
1926 *Quantum Mechanik der Stossvorgänge*, in «Zeitschrift für Physik», XXXVIII, p. 803.
1949 *Natural Philosophy of Cause and Chance*, Clarendon Press, Oxford (trad. it. Boringhieri, Torino 1962).

Brecher, K.
1977 *Is the speed of light independent of the velocity of the source?*, in «Physical Review Letters», XXXIX, pp. 1051-54.

Broglie, L.-V. de
1923 *Ondes et quanta*, in «Comptes Rendus de l'Académie de France», CLXXVII, p. 507.

Calladine, C. R.
1974 *Bacteria can swim without rotating flagellar filaments*, in «Nature», CCXLIX, pp. 385-86.

Cassirer, E.
1937 *Determinismus und Indeterminismus in der modernen Physik; historische und systematische Studien zum Kausalproblem*, Elanders, Göteborg (trad. it. La Nuova Italia, Firenze 1970).

Chwang, A. T., e Wu, T. Y.
1971 *A note on the helical movement of micro-organisms*, in «Proceedings of the Royal Society of London. Series B», CLXXVIII, pp. 327-46.

Copernico, N.
1543 *De revolutionibus orbium caelestium*, Petreium, Nürnberg (trad. it. Einaudi, Torino 1975).

Dirac, P.-A.-M.
1930 *The Principles of Quantum Mechanics*, Clarendon Press, Oxford (trad. it. Boringhieri, Torino 1959).

Eddington, A. S.
1928 *The Nature of the Physical World*, Cambridge University Press, Cambridge; ed. Macmillan, New York 1929.

Einstein, A.
1905*a* *Zur Elektrodynamik bewegter Körper*, in «Annalen der Physik», XVII, p. 891.
1905*b* *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt*, in «Annalen der Physik», XVII, p. 132.

Einstein, A., e Infeld, L.
1938 *The Evolution of Physics. The Growth of Ideas from Early Concepts to Relativity and Quanta*, Simon and Schuster, New York (trad. it. Boringhieri, Torino 1970[3]).

Galilei, G.
1638 *Discorsi e dimostrazioni matematiche intorno a due nuove scienze*, Elzevier, Leiden; ed. Boringhieri, Torino 1958.

Geymonat, L.
1973 *Federigo Enriques e la Storia della Scienza*, in *Atti del Convegno Internazionale «Storia, Pedagogia, Filosofia della Scienza», Roma 7-12 Ottobre 1971*, Accademia dei Lincei, quaderno n. 184, pp. 29-38.

Hamilton, W. R.
1834 *On a general method in dynamics*, in «Philosophical Transactions of the Royal Society of London», vol. CXXIV, p. 252.

Heisenberg, W.
1927 *Über der anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik*, in «Zeitschrift für Physik», XLIII, p. 172.
[1932] *Zur Geschichte der physikalischen Naturerklärung*, in *Wandlungen in den Grundlagen der Naturwissenschaft*, Hirzel, Stuttgart 1959[9] (trad. it. Boringhieri, Torino 1960, pp. 3-19).
[1948] *Die gegenwärtigen Grundprobleme der Atomphysik*, *ibid.* (trad. it. *ibid.*, pp. 138-55).

Hertel, H.
1963 *Struktur, Form, Bewegung*, Krausskopf-Verlag, Mainz.

Holwill, M. E. J.
1977 *Low Reynolds Number Ondulatory Propulsion in Organisms of Different Size-Scale Effects in Animal Locomotion*, Academic Press, New York.

Kármán, Th. von
1956 *Aerodynamics: Selected Topics in the Light of their Historical Development*, Cornell University Press, Ithaca N.Y. (trad. it. Interavia, Genève 1956).

Keplero, J.
1596 *Prodromus dissertationum cosmographicarum, continens mysterium cosmographicum, de admirabili proportione orbium coelestium...*, Gruppenbach, Tübingen; ora in *Gesammelte Werke*, vol. I, Beck, München 1938.

Knight-Jones, E. W.
1954 *Relations between metachronism and the direction of ciliary beat in Metazoa*, in «Quarterly Journal of Microscopical Science», XCV, pp. 503-21.

Lighthill, J.
1970 *Aquatic animal propulsion of high hydromechanical efficiency*, in «Journal of Fluid Mechanics», XLIV, pp. 265-301.
1971 *Large amplitude elongated-body theory of fish-locomotion*, in «Proceedings of the Royal Society. Series B», CLXXIX, pp. 125-38.

Neugebauer, O.
1975 *A History of Ancient Mathematical Astronomy*, Springer, Berlin - Heidelberg - New York.

Newton, I.
1713 *Philosophiae Naturalis Principia Mathematica*, Crownfield, Cambridge 1713[2] (trad. it. Utet, Torino 1965).

Planck, M.
1901 *Über das Gesetz der Energieverteilung in Normalspectrum*, in «Annalen der Physik», IV, 4, p. 553.

Rosen, M. M., e Cornfold, N. E.
1971 *Fluid friction of fish slimes*, in «Nature», CCXXXIV, pp. 49-51.

Schrödinger, E.
1926 *Quantisierung als Eigenwertproblem*, in «Annalen der Physik», LXXIX, pp. 361, 489; LXXX, p. 437; LXXXI, p. 109.

Siekmann, J.
1963 *On a pulsating jet from the end of a tube with application to the propulsion of certain aquatic animals*, in «Journal of Fluid Mechanics», XV, parte III, pp. 399-418.

Silverman, M., e Simon, M.
1974 *Flagellar rotation and the mechanism of bacteria motility*, in «Nature», CCXLIX, pp. 73-74.

Storch, O.
1929 *Die Schwimmbewegung der Copepoden auf Grund von Micro-Zeitlupenaufnahmen analysiert*, in «Verhandlungen der deutschen zoologischen Gesellschaft», XXXIII, pp. 118-29.

Taylor, G. I.
1952 *Analysis of the swimming of long and narrow animals*, in «Proceedings of the Royal Society of London. Series B», CCXIV, pp. 158-83.

Taylor, J. H.; Fowler, L. A.; e McCulloch, P. M.
1979 *Measurements of general relativistic effects in the binary pulsar PSR 1913-16*, in «Nature», CCLXXVII, pp. 437-40.

Vailati, G.
1897 *Il principio dei lavori virtuali da Aristotele a Erone di Alessandria*, in «Atti della Reale Accademia delle Scienze di Torino», XXXII (1896-97), pp. 940-62.

Weish, D.
1973 *The mechanism of rapid starting of slender fish*, in «Biorheology», X, pp. 343-50.

Wichmann, E. H.
1967 *Quantum Physics*, McGraw-Hill, New York (trad. it. Zanichelli, Bologna 1978).

Wu, T. Y.
1971 *Hydromechanics of swimming propulsion*, in «Journal of Fluid Mechanics», XLVI, pp. 337-55, 365-68.

Tutto nella **natura** si muove e il movimento è la prima indicazione della **vita**: si comprende quindi come il termine 'moto' dal primo significato di cambiamento della situazione spaziale di un oggetto sia venuto a indicare ogni mutazione interessante anche gli attributi qualitativi (cfr. **qualità/quantità**) di un dato individuo e in particolare le forme di mutazione (cfr. **forma, sviluppo e morfogenesi**) sia «per salti» sia «continue» (cfr. **continuo/discreto**), il piú delle volte misurabili (cfr. **misura**) nel tempo (cfr. **tempo/temporalità, ciclo, evento**). Se dal punto di vista delle filosofie (cfr. **filosofia/filosofie**, e anche **dialettica, metafisica**) il problema del movimento-mutamento può venir ricondotto alla questione dell'**identità/differenza**, per le **matematiche** l'intuizione del moto ha sotteso lo sviluppo del **calcolo** (cfr. anche **infinitesimale, differenziale, curve e superfici, funzioni, variazione** e, per altri aspetti, **sistemi di riferimento, invariante, geometria e topologia, stabilità/instabilità, catastrofi**). Già lo studio delle proprietà meccaniche della **materia**, con la costituzione di cinematica, statica e, soprattutto, dinamica ha fornito alla **fisica** un **paradigma** di scientificità (cfr. **teoria/modello, scienza, conoscenza**) capace di unificare (cfr. **uno/molti**) i fenomeni (cfr. **fenomeno**) piú disparati dell'**universo**, come i moti terrestri e celesti (cfr. **astronomia, cosmologie** e soprattutto **gravitazione**). L'evoluzione di concetti fisici primari come quelli di spazio e di tempo (cfr. **spazio-tempo**), l'arricchimento del patrimonio teorico e sperimentale della fisica (cfr. in particolare **energia, entropia, reversibilità/irreversibilità**), le concettualizzazioni in termini di **forza/campo**, ecc. hanno infine riproposto, specie nel contesto relativistico (cfr. **relatività**) e soprattutto in quello quantistico (cfr. **quanti**, e anche **particella**), la questione centrale della spiegazione causale (cfr. **causa/effetto**) e quella, connessa, del determinismo fisico (cfr. **determinato/indeterminato**, e anche **caso/probabilità**).

Ma i moti naturali negli esseri viventi, in particolare negli animali (cfr. **animale**, e anche **homo**) pongono oggi problemi altrettanto affascinanti: tali moti obbediscono ovviamente alle leggi della fisica, ma le modalità con cui possono venir realizzati non sono univocamente determinate dall'**ambiente** in cui vive la **specie** considerata. Tali modalità soddisfano inoltre vincoli (cfr. **vincolo**) dipendenti, non solo dall'ambiente, ma anche dalle caratteristiche strutturali dei vari organismi (cfr. **organismo**) o da altre circostanze. L'ottimizzazione che le specie animali riescono a realizzare nella locomozione è, in genere, attribuita a un meccanismo di **mutazione/selezione**, il che ripropone nuovamente il problema di una natura che agisce a caso o seguendo un **programma**.

Nota alle tavole.

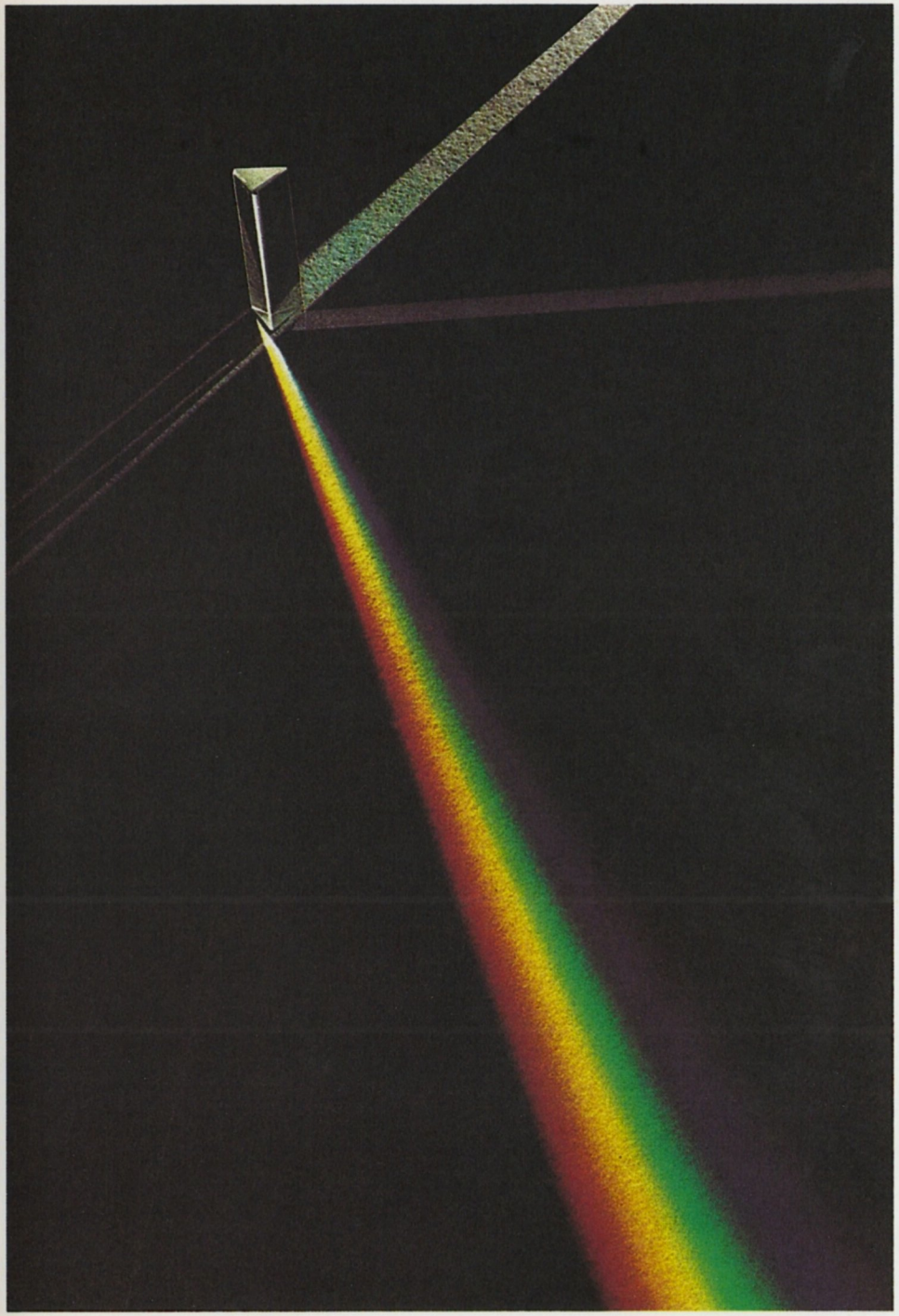
«SOCRATE. E invece... di chiaro apparisce questo, che, se tutto si muove, ogni risposta, su qualunque cosa uno risponda, è egualmente giusta, sia che si dica che la cosa "sta cosí", sia che "non sta cosí"... TEODORO. Dici bene. SOCRATE. Sí, Teodoro; tranne che dissi 'cosí' e 'non cosí'. E invece neanche questo 'cosí' si deve dire, perché non ci sarebbe piú moto dicendo di una cosa 'cosí'; e nemmeno 'non cosí', perché neppure questo è moto; ma bisogna che istituiscano un altro linguaggio coloro che professano questa dottrina» [Platone, *Teeteto*, 183a-b). Ma se per la conoscenza (si vedano del resto in questa stessa *Enciclopedia* gli articoli «Conoscenza», «Filosofia/filosofie» e «Dialettica») la dottrina del moto fa esplodere delle contraddizioni, una sfida non minore costituisce la sua *rappresentazione*. Come è possibile rendere con una combinatoria discreta di forme quel fluire continuo che il Socrate platonico riconosce intrinseco al movimento? Già una prima risposta è nella modellizzazione astronomica dove è la stessa immaginazione geometrica a suggerire significative visualizzazioni (tavv. 10-13; cfr. del resto quanto osservato nell'articolo «Modello» in questa stessa *Enciclopedia*): appartenga alla tradizione orientale (cfr. in particolare tav. 13) o a quella occidentale (cfr. in particolare tav. 10), segni ai tempi di Copernico, Keplero o Galileo (cfr. in particolare tavv. 11-12) una rottura «rivoluzionaria» o possa venir letta, con gli occhi di oggi, come anticipazione di attuali cosmologie (cfr. ancora tav. 13), questo tipo di visualizzazione fornirà un paradigma anche per altre discipline dovunque esse hanno a che fare col moto.

Alcune di queste tavole alludono quindi alle ancor piú sofisticate modellizzazioni emerse in tali discipline: dalla natura e propagazione della luce (cui allude la tavola 1: non è del resto la luce, come già diceva Roberto di Lincoln, a rivelarci il movimento dei corpi?) a complesse situazioni idrodinamiche e aerodinamiche (cfr. tav. 4), dal trapasso delle stagioni e dei movimenti nell'atmosfera (cfr. tav. 5) fino allo studio dei movimenti/mutamenti dell'idrosfera e della litosfera (cfr. tavv. 21 e 22, ma anche tav. 19).

Spesse volte la rappresentazione ha preceduto modellizzazioni articolate concettualmente e linguisticamente, altre volte ne è l'indispensabile ausilio. Cosí è, per esempio, nello studio del movimento animale: sia che si tratti della figura di un cervo schizzata sulla parete di una delle grotte di Lascaux (tav. 2), sia che si tratti di un animale in corsa colto da una sequenza fotografica (tav. 3), la rappresentazione coglie quegli aspetti «meravigliosi» propri del moto animale che ancora oggi stupiscono per la genialità della soluzione che naturalmente essi hanno avuto, e inserisce questa «meraviglia» nei contesti piú vari: da quelli incentrati sulla relazione predatore/preda (come la caccia, ecc.), a quelli che costituiscono articolazioni della coppia sacro/profano, ai contesti piú propriamente scientifici o piú nettamente tecnologici.

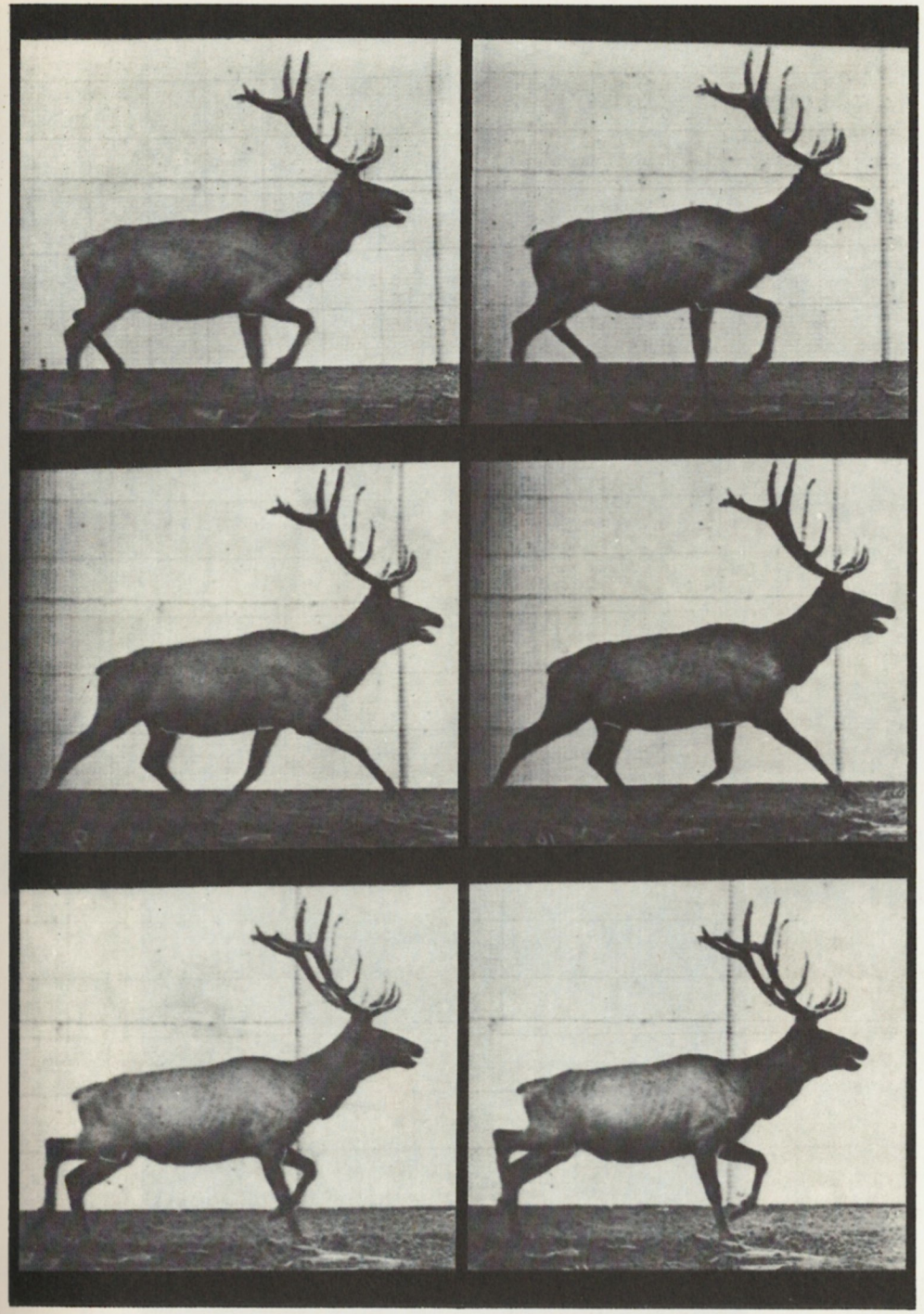
Dal moto dei vari animali (cfr. ancora tav. 23, e anche le numerose figure inserite nel testo dell'articolo), infine, al moto dell'«animale» uomo: il gesto (tav. 29), la corsa (tav. 15), l'attività sportiva (tav. 14), ecc. diventano oggetti, a loro volta, di rappresentazione e modellizzazione (cfr. tav. 16). Né manca l'imitazione tecnologica: il volo di Icaro (tav. 6) prelude, al di là del mito, al volo aereo, prima sognato, poi audacemente tentato (tav. 7), e infine realizzato (tavv. 8-9). Del resto l'attuale relazione dell'uomo col suo ambiente sarebbe incomprensibile senza tener conto delle realizzazioni delle potenzialità meccaniche che l'uomo ha vagheggiato fin dall'aurora del pensiero scientifico e che in tempi recenti hanno radicalmente trasformato i mezzi aerei (cfr. ancora tav. 9), acquatici (cfr. tav. 24) e terrestri (cfr. tavv. 17, 28; 25, 26, 27; e anche 30).

Ma agli strumenti tecnologici che sempre di piú permettono all'uomo lo sfruttamento del moto, si affiancano (e s'intrecciano) gli strumenti intellettuali dell'immaginazione e dell'astrazione: si risale cosí all'inizio del mondo (cfr. ancora tav. 13) o ci si proietta verso enigmatiche forme del futuro (tavv. 32-33), cercando nuove risposte al semplice 'cosí' o 'non cosí' (cfr. anche tavv. 18, 20). [*N.d.E.*].

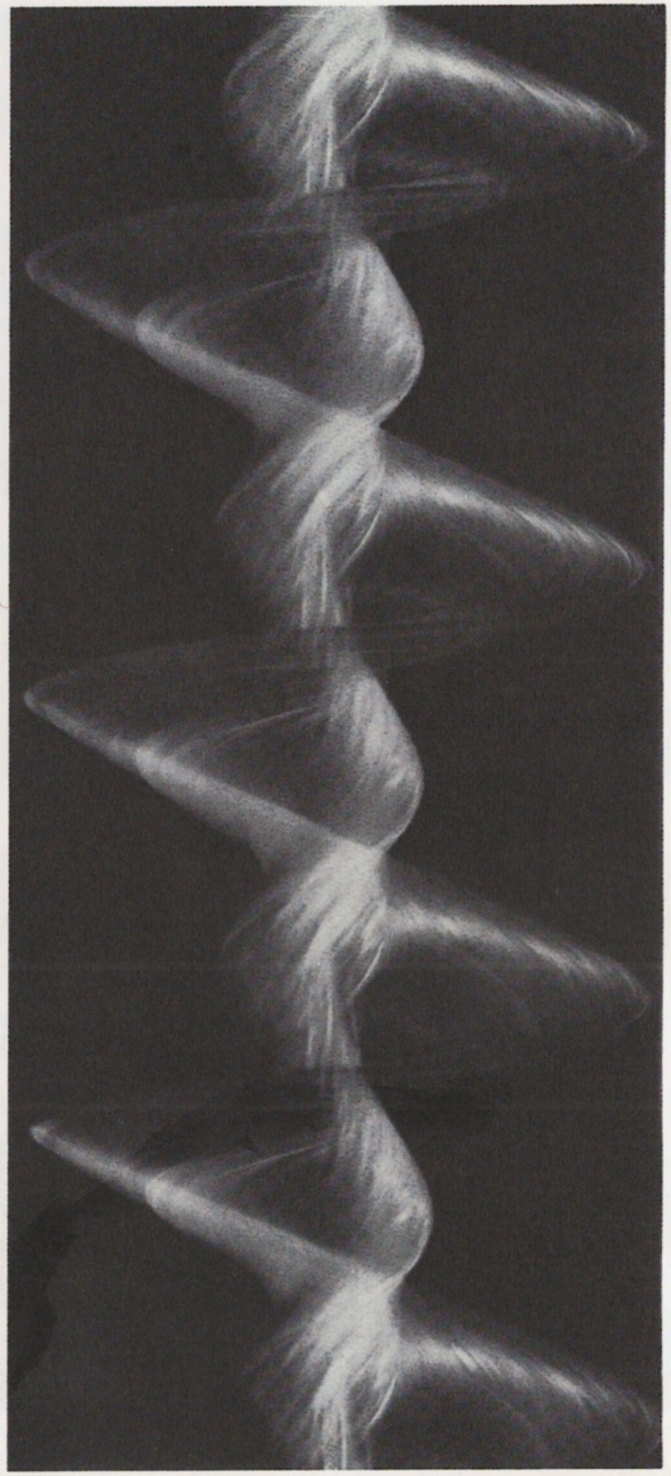


1. *Prism and Spectrum.* Foto di Sebastian Milito (1970).

2. Teste di cervo. Graffito delle grotte di Lascaux.



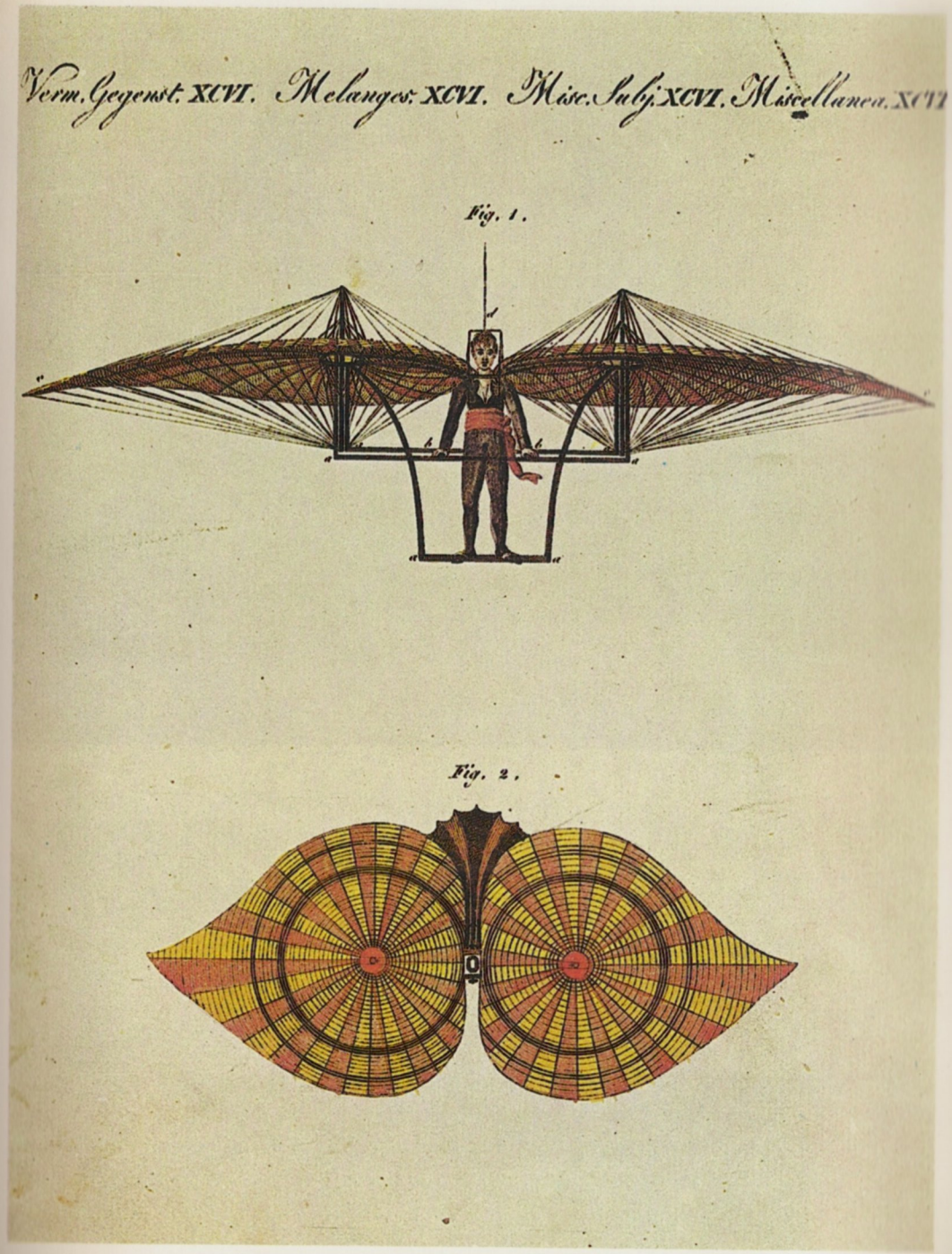
3. Alce che corre. Parte di una sequenza fotografica di E. Muybridge (1887 circa).



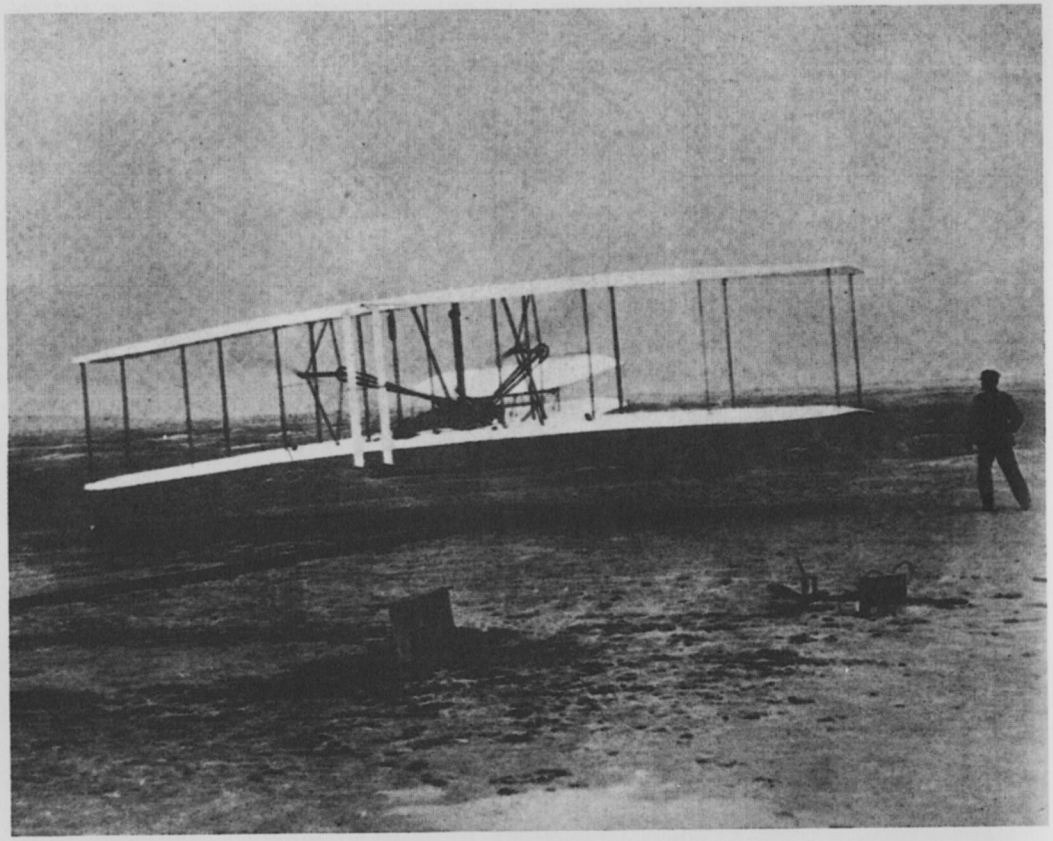
4. Frutto di *Hiptage madablota* in caduta libera.

5. Ch'ien Hsüan, *Primo autunno* (secolo XIII). Particolare.

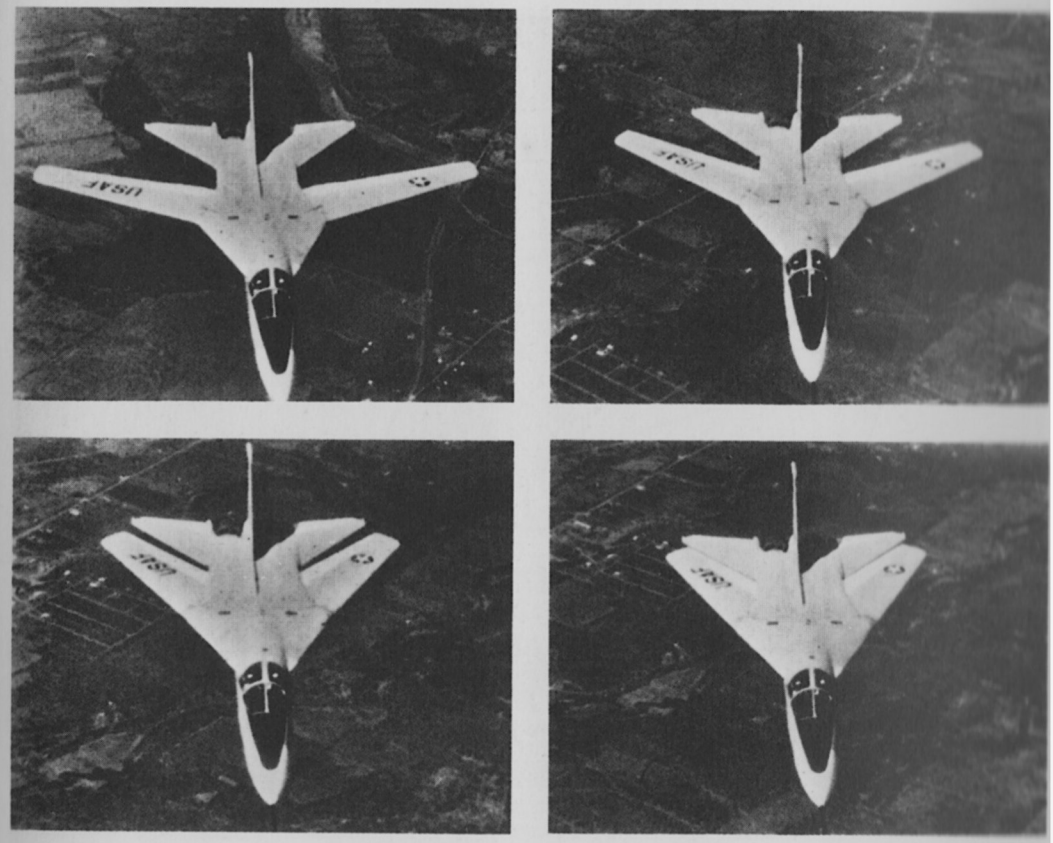
6. *Icaro* di Célestin Nanteuil (secolo XIX).



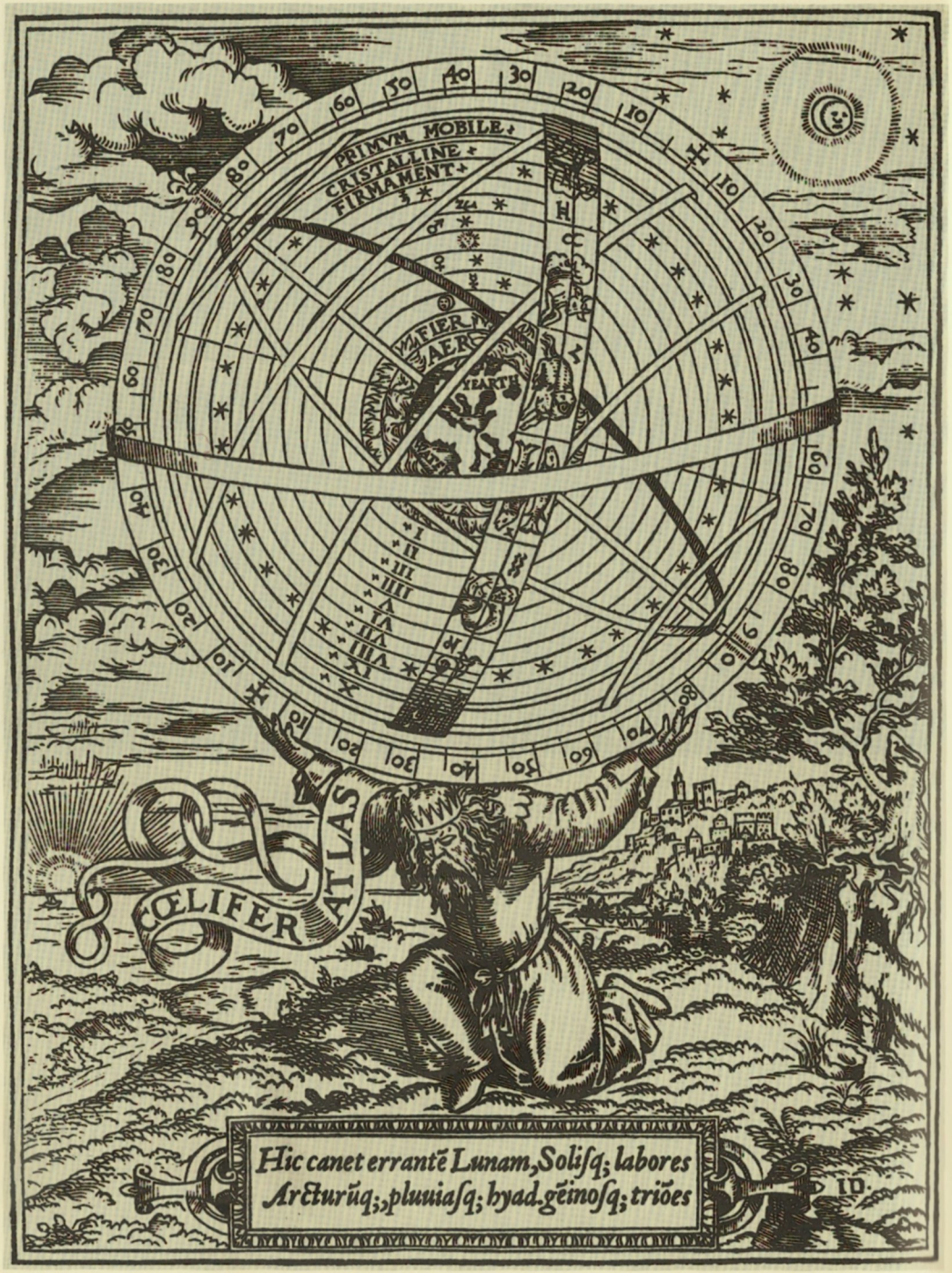
7. Tentativo di Jacob Dengen di sollevarsi in aria mediante un apparecchio di sua invenzione (1808).



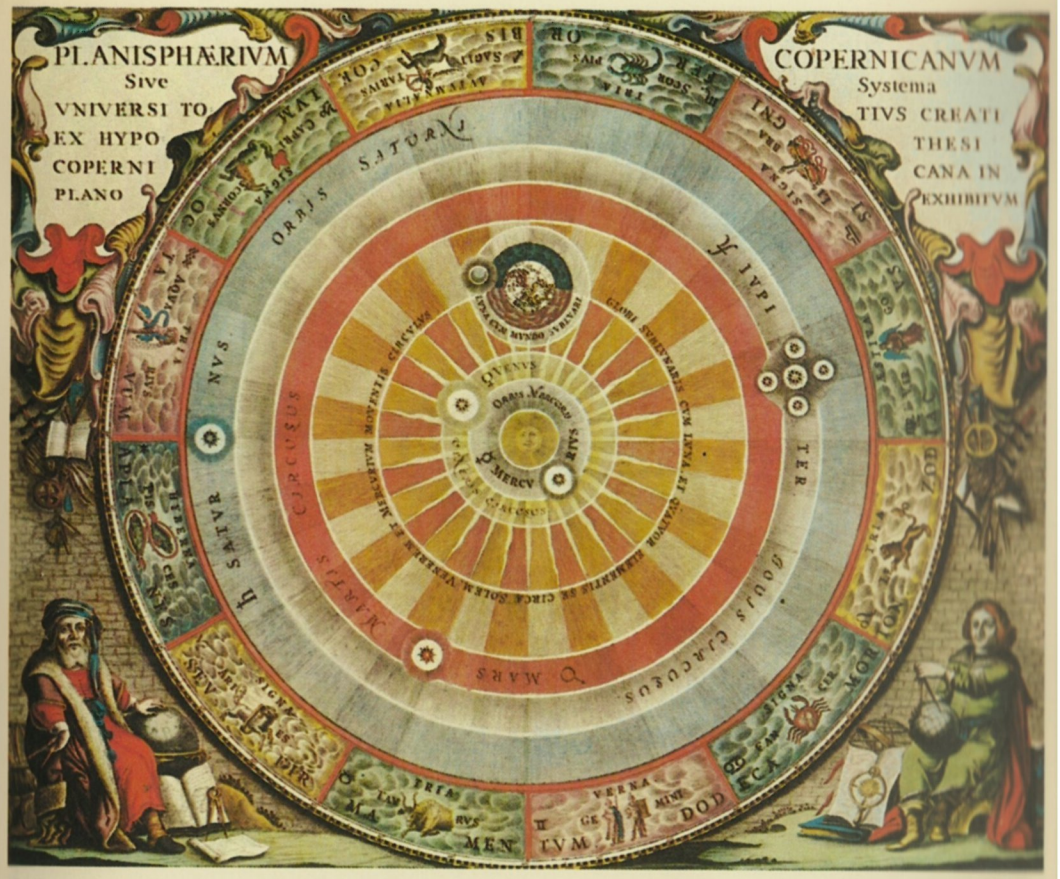
8. Orville Wright vola per la prima volta sul suo apparecchio (17 dicembre 1903). Foto di John D. Daniels.



9. Il bireattore F-111 con ali a freccia variabile.



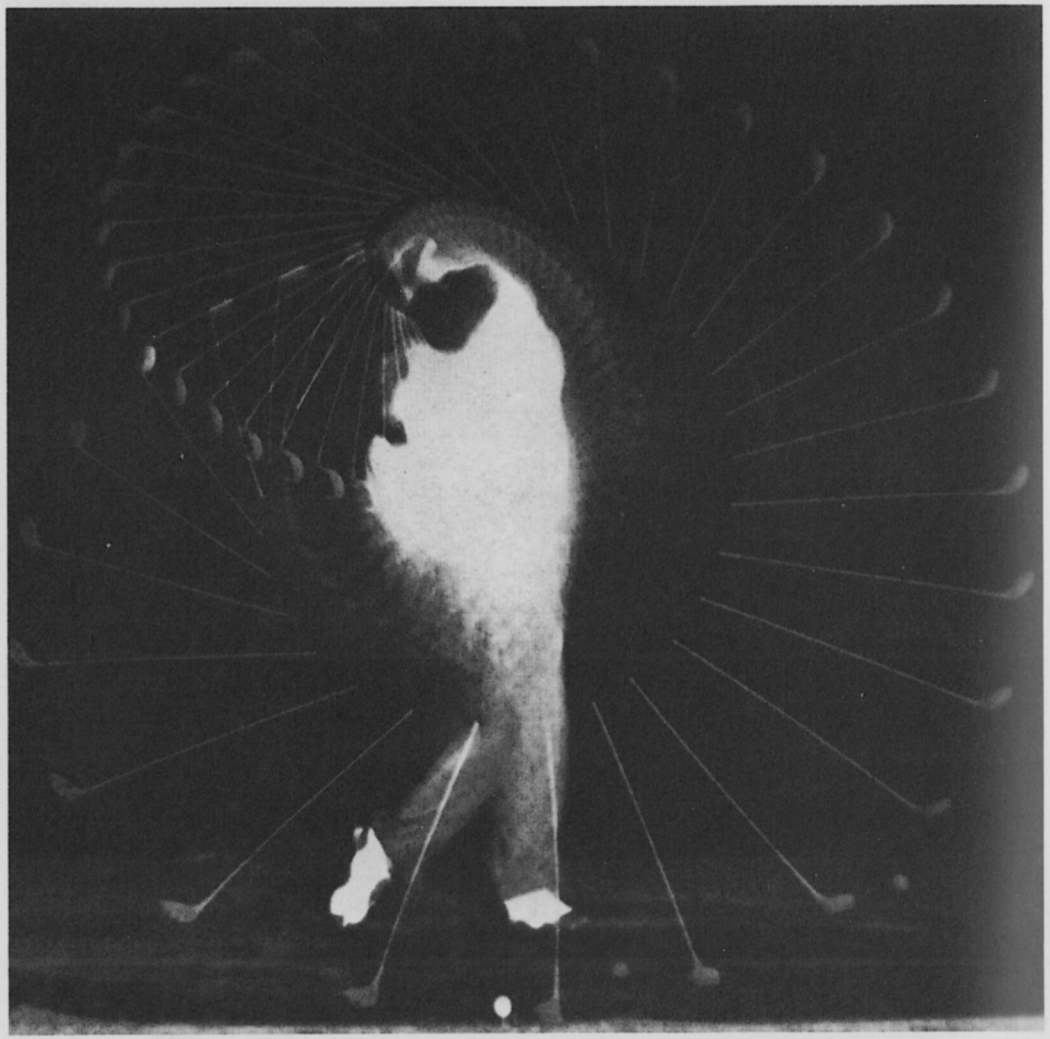
10. Atlante sorregge il mondo: al centro, la Terra (1559).



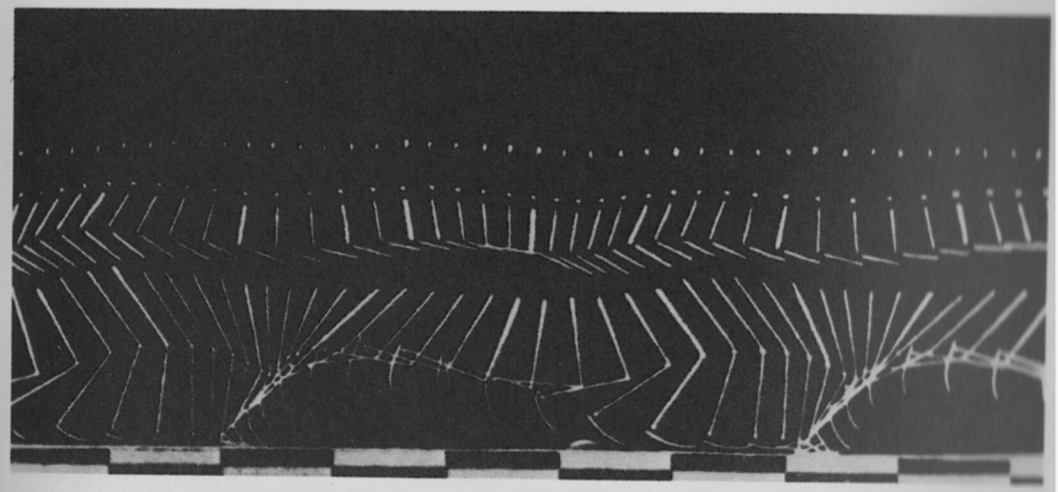
11. Il sistema eliocentrico copernicano. Dall'*Atlas coelestis* di A. Cellarius (1660).

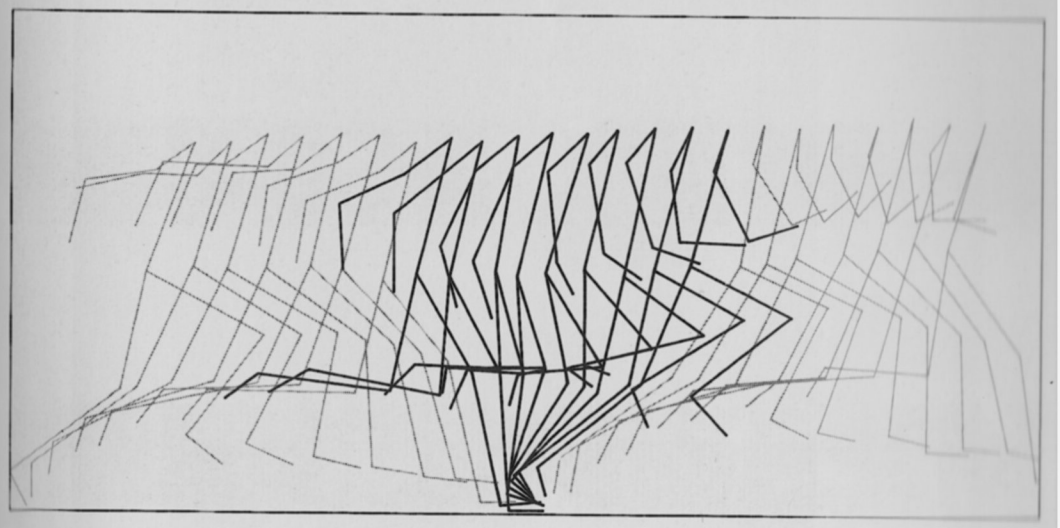
12. Figure che accompagnano il primo abbozzo del *Sidereus nuncius*, autografo di G. Galilei (circa 1610).

13. L'universo in espansione (Rājasthān, secolo XVIII).



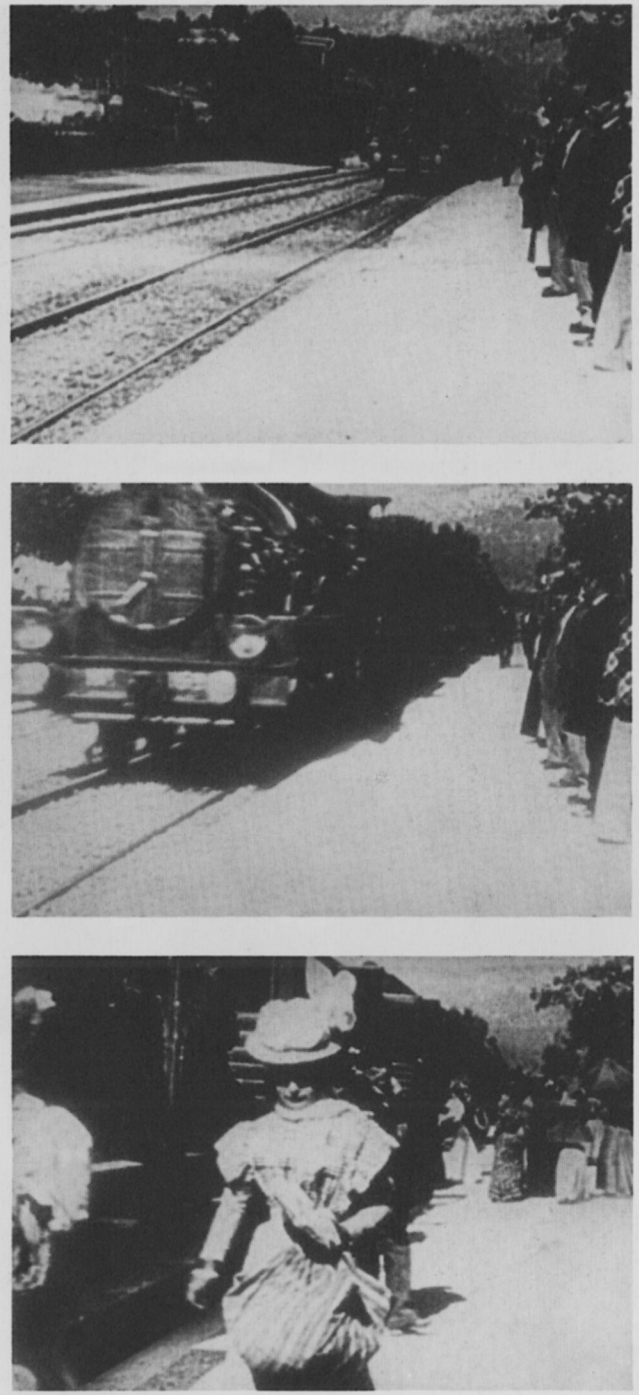
14. Movimento del golfista D. Shute fotografato da Harold E. Edgerton (1938).



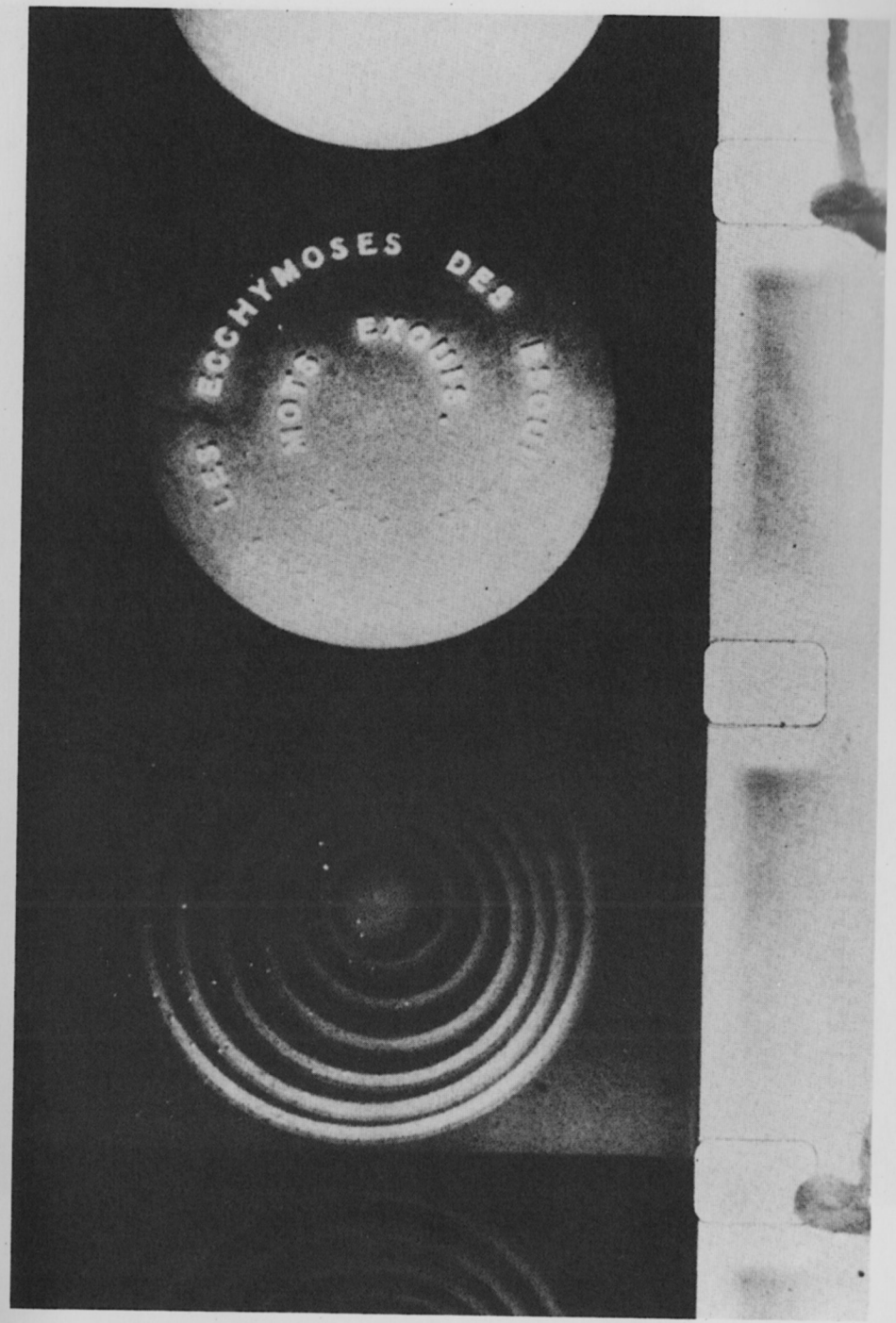


15. Uomo in corsa (circa 1883). Studio fotocronografico di E.-J. Marey.

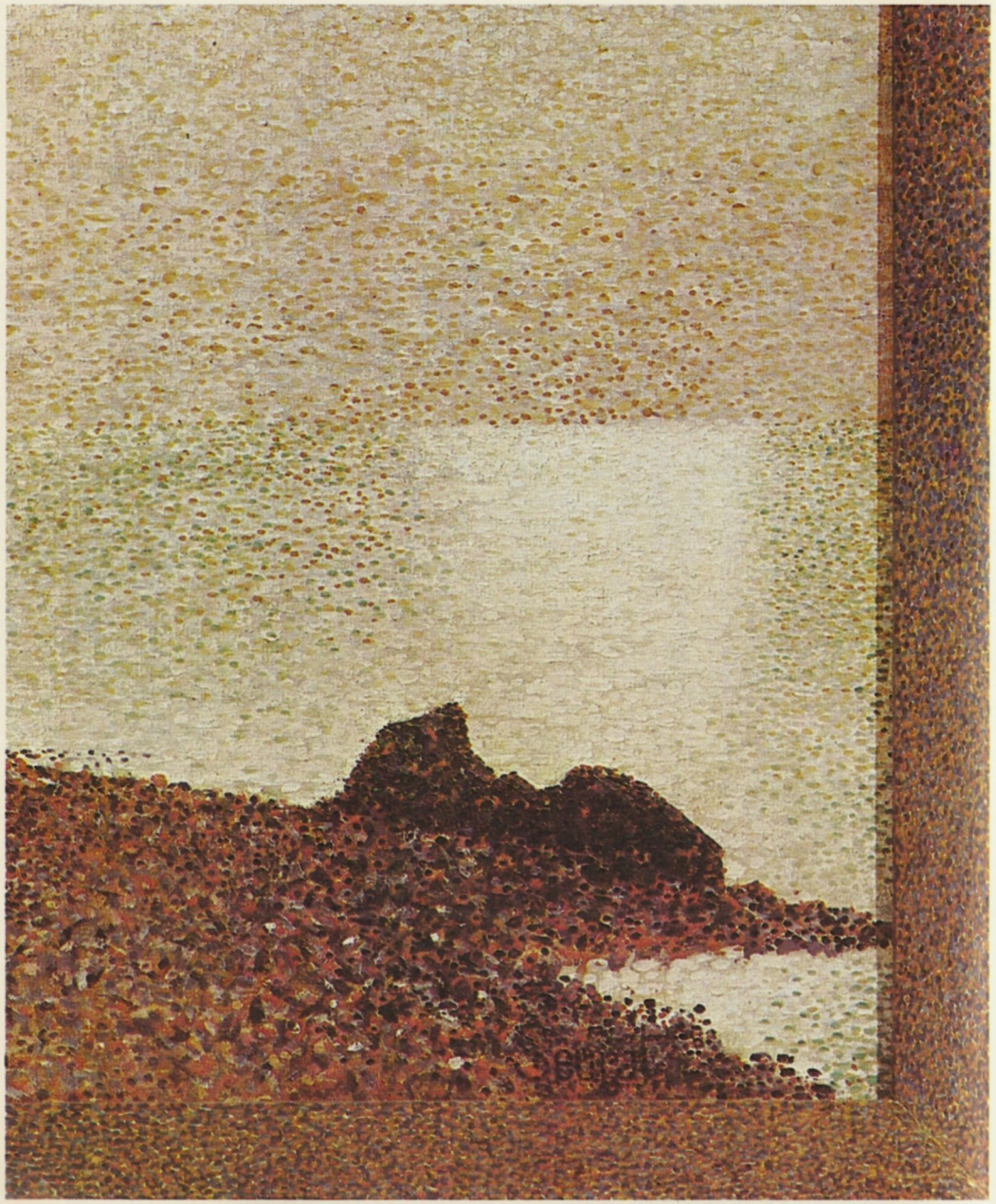
16. Movimento di un atleta in corsa (1976) analizzato da un calcolatore elettronico.



17. Da *L'arrivée d'un train* di Louis Lumière (1895).

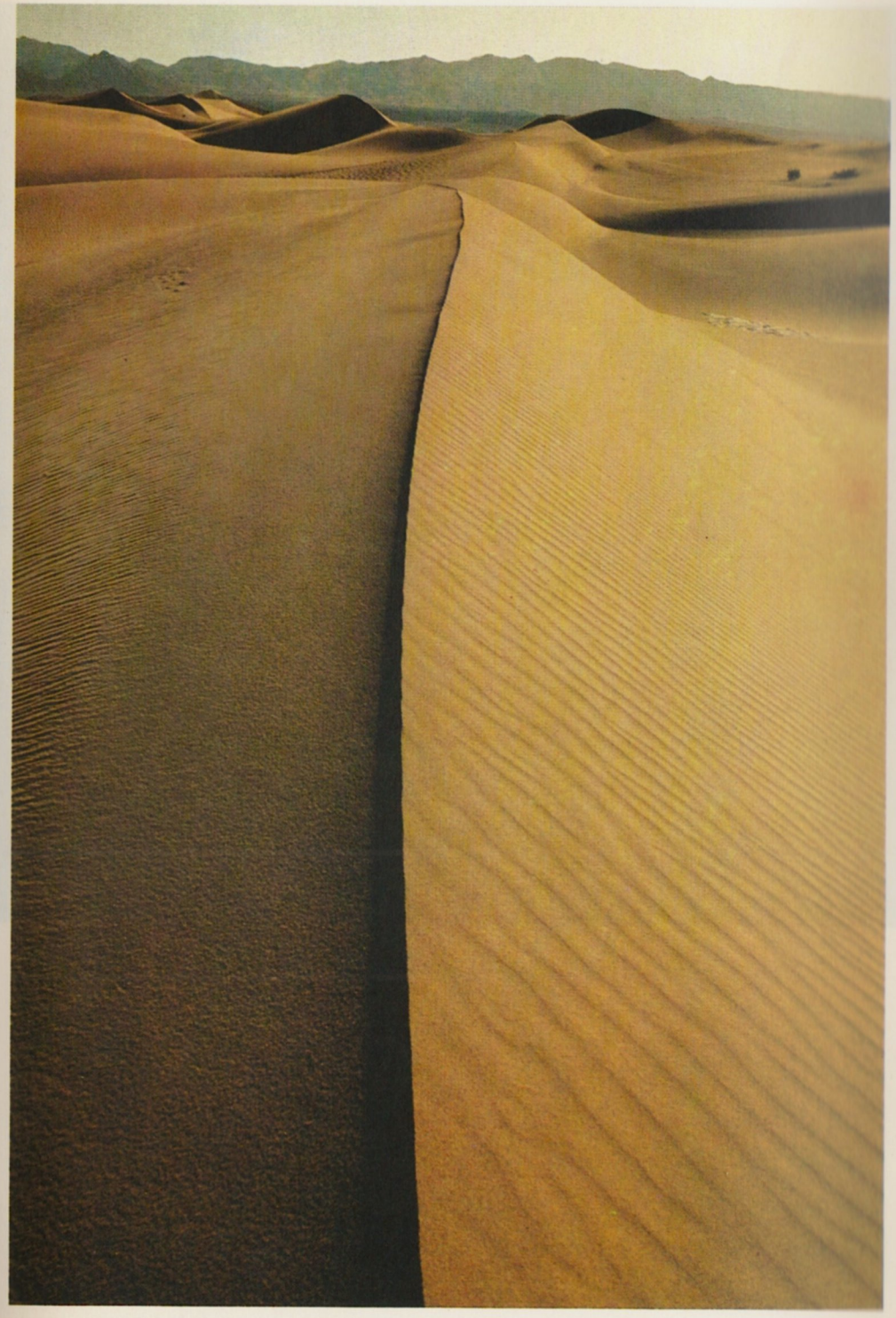


18. Marcel Duchamp, fotogrammi del film *Anemic Cinema* (in collab. con Man Ray e Marc Allégret) (1925-26).



19. *La foce della Senna a Honfleur, di sera* di G. Seurat (1886). Particolare.

20. *Forme uniche della continuità dello spazio* di U. Boccioni (1913).

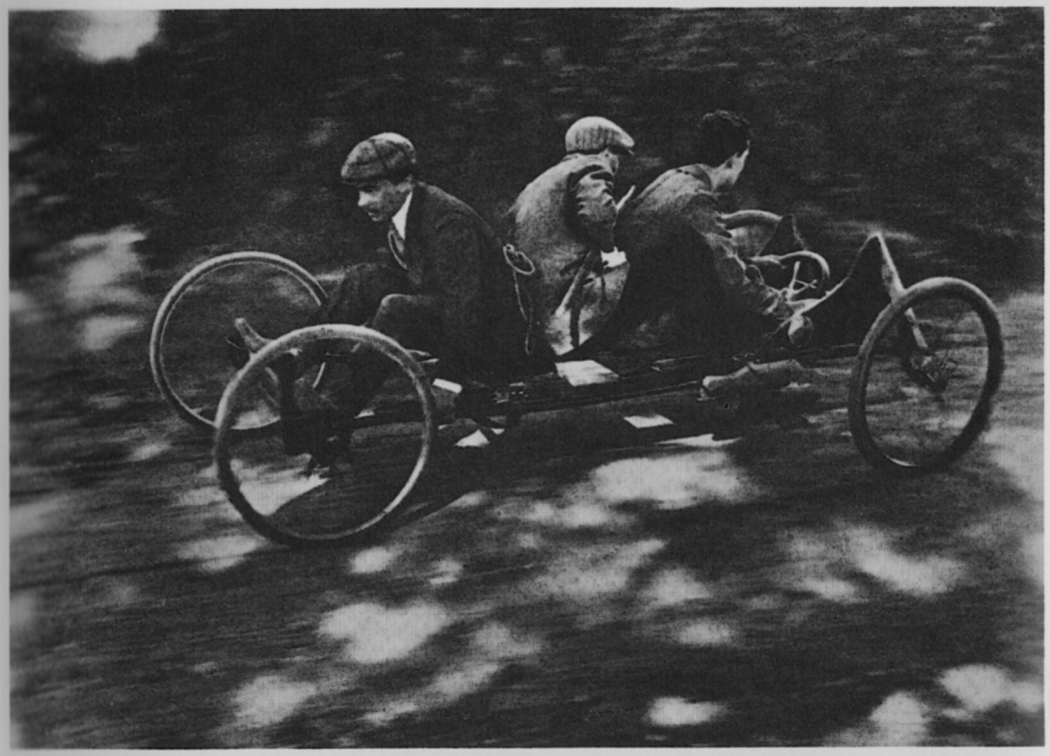


21. Il ghiacciaio Barnard. Foto di Bradford Washburn (1938).

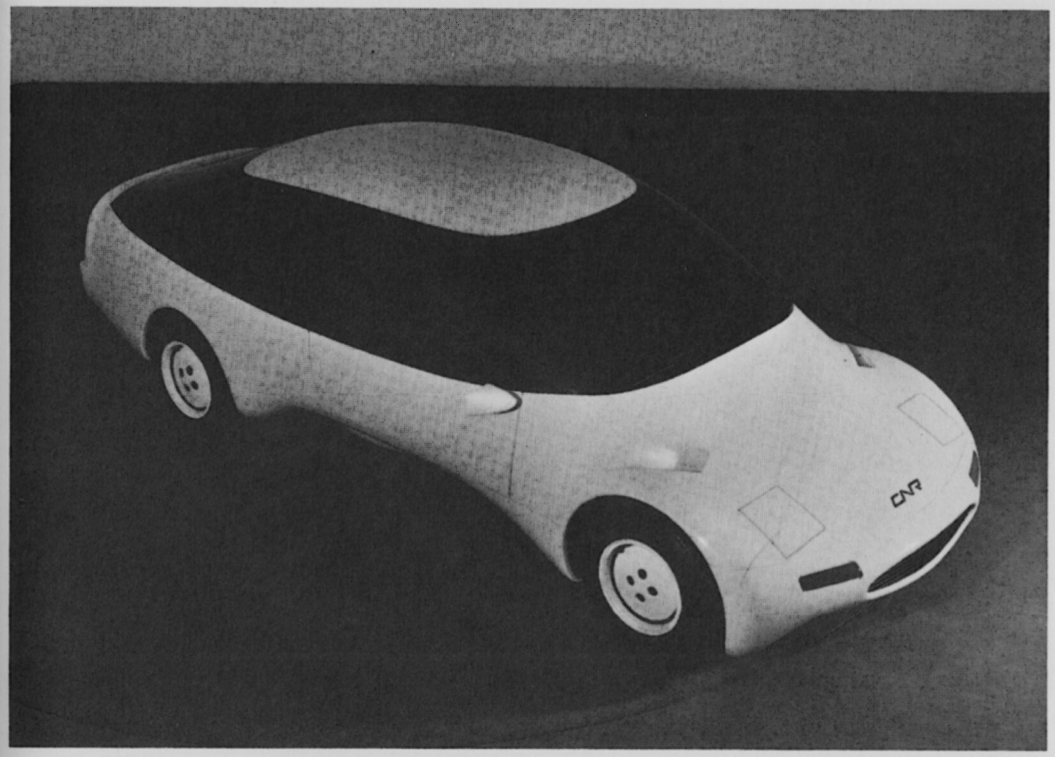
22. Death Valley. Foto di Charles Moore (1970).

23. Granchio d'acqua. Foto di John Cooke (1969).

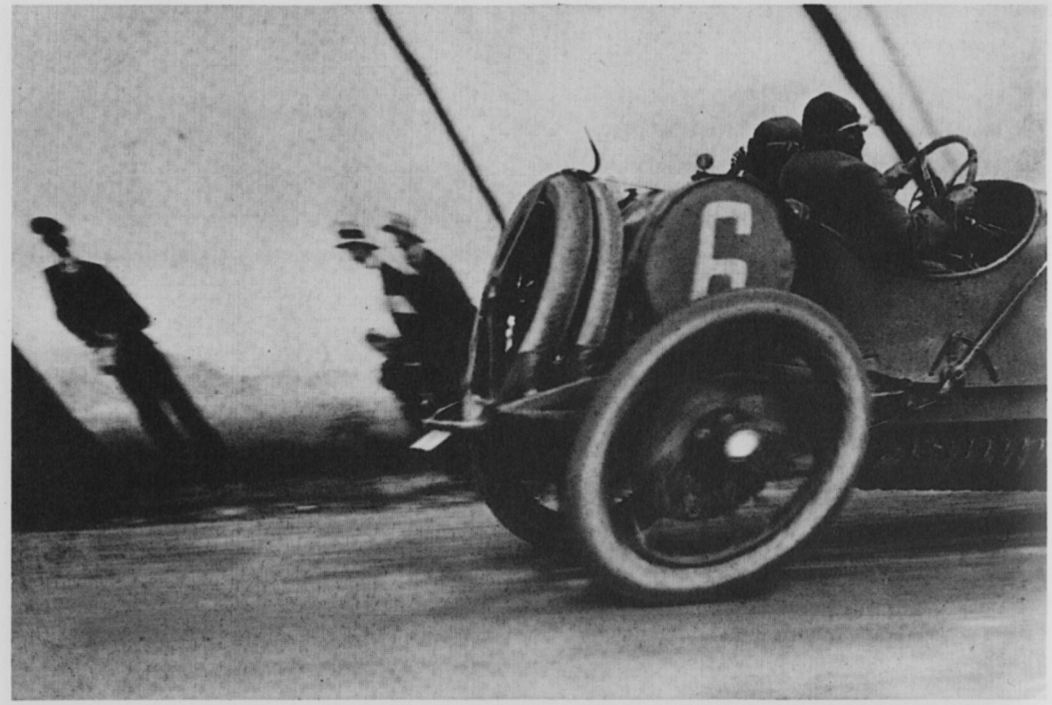
24. La fregata a vapore Fulton II (circa 1850). Particolare.



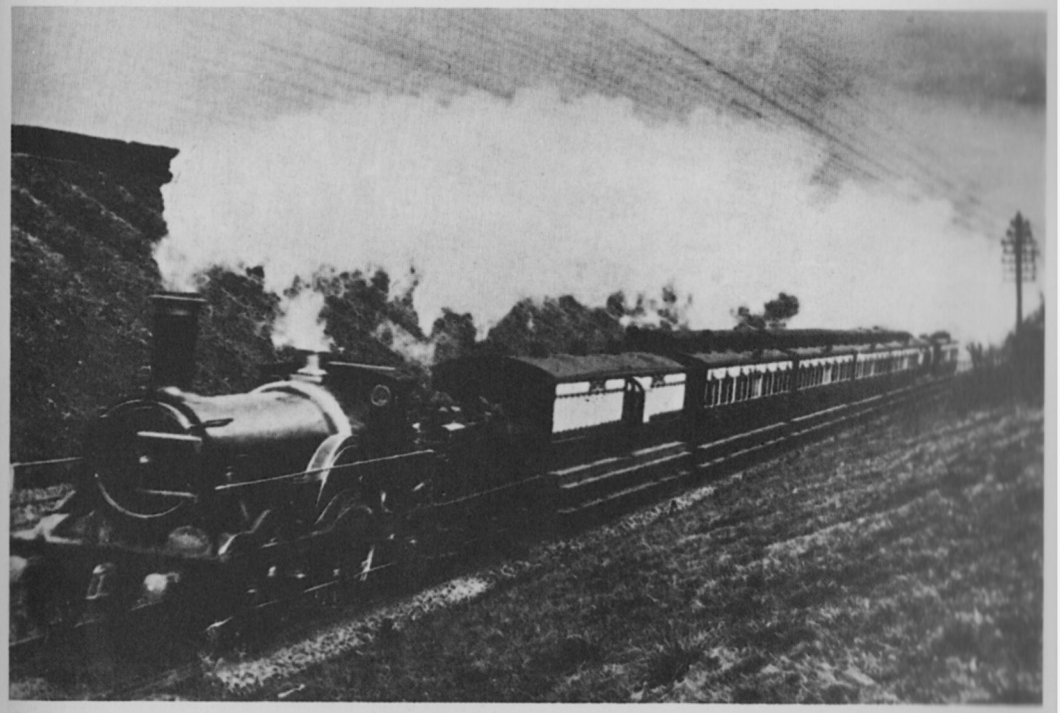
25. *Course de bobs. Louis-Oleo-Zissou* (1910). Foto di Jacques-Henri Lartigue.



26. Studio di forma aerodinamica ideale (1979).



27. Una Delage del 1912. Foto di Jacques-Henri Lartigue.



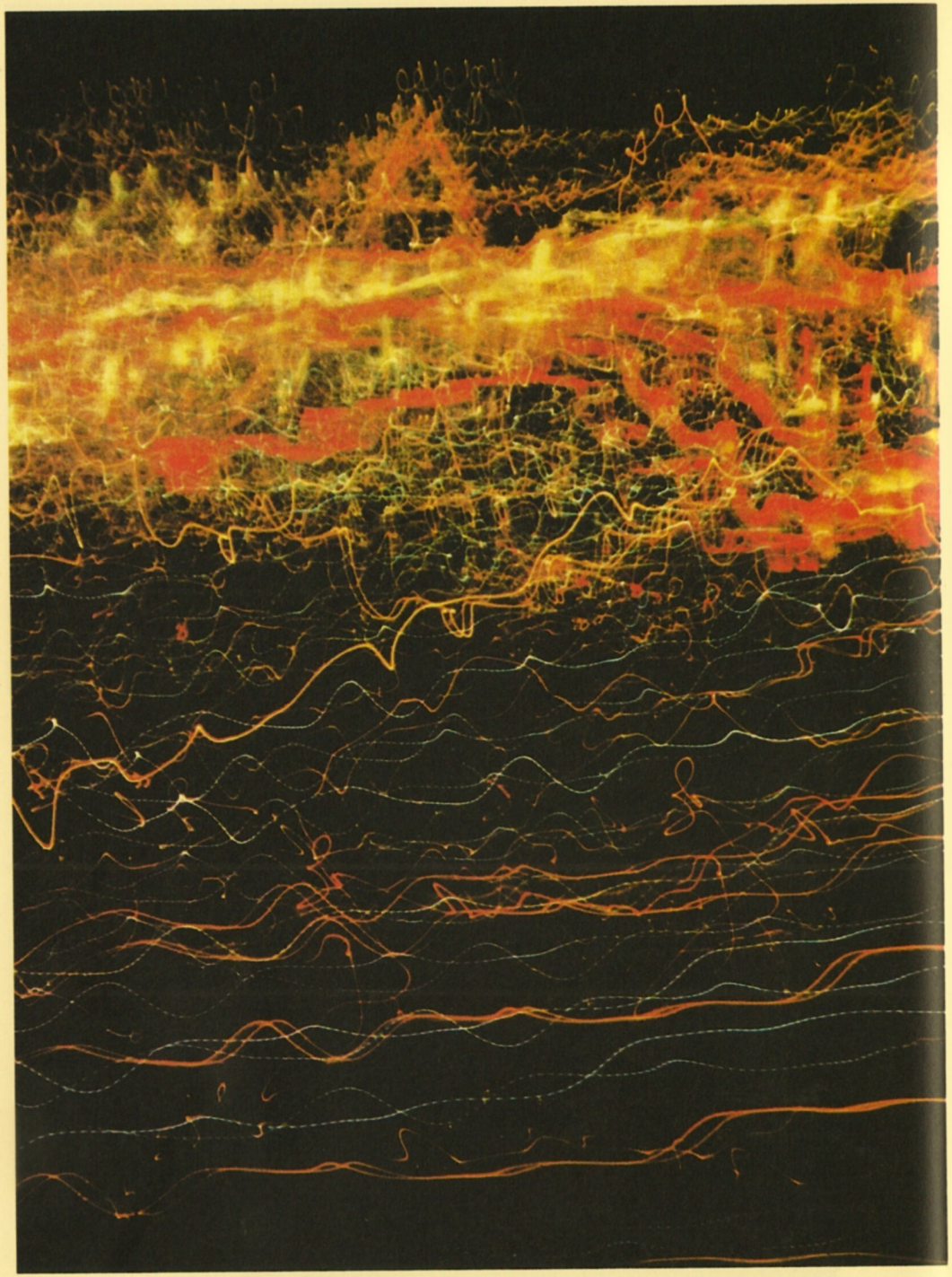
28. L'espresso Londra-Brighton (1875).

29. Una donna porge una giara alla compagna (circa 1887). Sequenza fotografica di E. Muybridge.

30. Moto in corsa (1931). Foto di Martin Munkacsy.

31. *Il Giudizio Finale* di Jacopo Robusti detto il Tintoretto (1560). Particolare.

32. *Liberazione* di M. C. Escher (1955).



33. *Aerial Abstract* (1969). Foto di Robert Mayer.

Particella

1. *Preliminari. I «punti materiali» e le «particelle elementari».*

1.1. Un esperimento mentale.

S'immagini un professore di fisica della metà dell'Ottocento che insegna la meccanica ai suoi allievi del primo anno. Poiché è in ritardo sul programma, decide di raddoppiare le sue dimostrazioni, accoppiandole e facendole seguire l'una all'altra. La cosa sembra funzionare: il professore arriva cosí all'esperimento del piano inclinato di Galileo e lo raddoppia nel modo seguente. Anziché terminare la loro corsa alla base del piano inclinato, le sfere vengono fatte cadere in un labirinto attraverso i cui chiodi e deflettori esse si distribuiscono. Ogni sfera cade infine in uno dei contenitori tubolari della figura 1. Naturalmente il professore si aspetta che le palle si raccolgano insieme come nella figura 2*a*. Se cosí avvenisse, egli potrebbe spiegare la caduta libera dei corpi e le distribuzioni statistiche tutto in un'ora di lezione. Inspiegabilmente, le sfere si distribuiscono invece nel modo illustrato nella figura 2*b*: una distribuzione improbabile, che non ha precedenti nella teoria classica dei punti materiali.

L'ipotetico professore non potrebbe dare del fenomeno che una sola spiegazione in accordo con il sapere della sua epoca: la distribuzione «anomala» ottenuta costituisce un tipico caso di interferenza, il segno, per cosí dire, di un processo periodico, ondulatorio. Nella schematizzazione della figura 1 la parte superiore definisce un comportamento corpuscolare delle palle; la parte inferiore, invece, un loro comportamento ondulatorio. Per descrivere queste scoperte in modo adeguato il professore dovrebbe combinare in *un unico* linguaggio *due* notazioni, quella corpuscolare e quella ondulatoria.

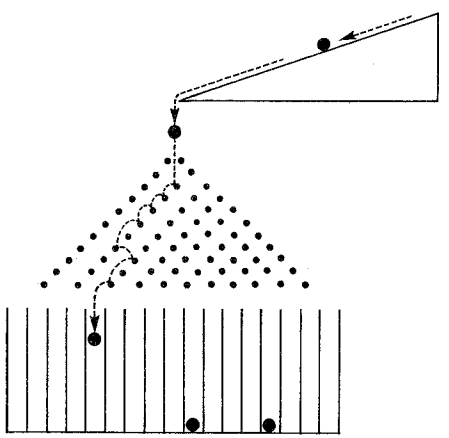


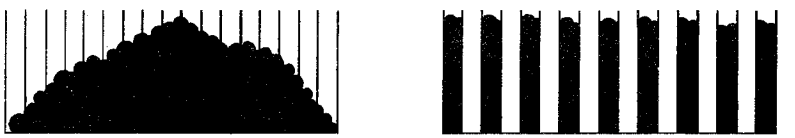
Figura 1.

L'accoppiamento di due esperienze (caduta di un grave lungo un piano inclinato e distribuzione casuale delle palline in una serie di contenitori attraverso un labirinto). (Da Hanson 1958).

L'esperimento mentale che qui si è brevemente delineato [sulla scorta di Hanson 1958, trad. it. pp. 150-51] è una parabola della fisica novecentesca. L'intera teoria classica dei punti materiali e l'intera teoria ondulatoria – che il lettore troverà entrambe esposte nel § 1 dell'articolo «Moto» in questa stessa *Enciclopedia* – sono infatti coinvolte nel riconoscimento da parte dell'ipotetico professore di essersi imbattuto in un fenomeno per cui si richiede un approccio combinato. Ed è proprio questo ciò che fece Einstein con la luce [1905]. «Poiché la luce proietta ombre nette e non s'incurva attorno agli angoli, Newton paragonò la luce a un flusso di particelle. Huygens, Young, Fresnel e Foucault screditarono quest'opinione, formulando le loro classiche idee della luce lungo linee ondulatorie. Einstein eccepí, spiegando l'effetto fotoelettrico mediante quanti di radiazione [fotoni] modellati sulle ricerche di Planck. Benché gli esperimenti di J. J. Thomson avessero indotto Lorentz a introdurre la discontinuità nella teoria di Maxwell, la contraddittorietà di questo intervento non emerse finché Einstein non uní esplicitamente proprietà corpuscolari e ondulatorie nel suo saggio del 1905» [Hanson 1958, trad. it. p. 239]. I fisici del tempo consideravano la *natura duale* della luce un paradosso. Ma è storia nota: questo «fenomeno sorprendente» fu la molla principale della teoria quantistica, almeno in una prima sua fase (si veda del resto l'articolo «Luce» in questa stessa *Enciclopedia*).

1.2. Onde e corpuscoli.

Il punto di partenza della teoria sviluppata da Louis de Broglie a partire dal 1923 (si veda ancora l'articolo «Moto», vol. IX, pp. 606 sgg.) si è basato sull'approccio onda-corpuscolo seguito da Einstein per il fotone, per darne un'estensione all'elettrone. Per Schrödinger (1926; si veda ancora «Moto», pp. 609-13) il moto dell'elettrone è descritto da un'onda associata, la funzione d'onda, soluzione di un'equazione differenziale alle derivate parziali (equazione di Schrödinger). Come lo stesso Schrödinger ebbe a mostrare, il formalismo ondulatorio e quello matriciale della «meccanica quantistica» sono, sotto un certo profilo, equivalenti. La ψ di Schrödinger, infine, secondo l'interpretazione datane da



a) *b*)

Figura 2.

La distribuzione statistica attesa (*a*) e quella «anomala» (*b*). (Da Hanson 1958).

Born, rappresenta lo stato fisico di un sistema, permettendo il calcolo dei valori medi delle grandezze fisiche che descrivono il sistema in quello stato dato.

Come il lettore sa da altri articoli della presente *Enciclopedia*, è stata questa una delle piú rilevanti svolte della fisica novecentesca: fino ad allora «la teoria quantistica... fu un aggregato di frammenti sostanzialmente diversi... e parzialmente contraddittori, [come per esempio] il principio di corrispondenza, appartenente per metà alla meccanica classica e all'elettrodinamica [e] l'autocontraddittoria natura duale della luce» [Neumann 1932, p. 4]. Questo punto merita di venir approfondito.

Il termine 'particella', di per sé, non definisce un sistema fisico: nell'immensità dell'universo, la Terra o il Sole sono particelle; nelle dimensioni macroscopiche della vita quotidiana un granello di polvere è una particella; nell'intima struttura della materia sono particelle atomi e molecole; entro l'atomo lo sono, infine, i nucleoni, gli elettroni e i partoni – dato e non concesso che questi ultimi esistano... Alcune di queste accezioni possono venir riportate a quello che è il concetto classico di particella, cioè di punto privo di dimensioni ma dotato di massa che, almeno in linea di principio, ha, in ogni tempo, una esatta posizione spaziale e un'esatta velocità. Alcune, ma non tutte.

Dalla proposizione che P è una particella (nel senso della meccanica classica poco piú sopra chiarito) si può certo concludere che P ha precise e determinabili posizione e velocità. «Designiamo questa proposizione come "Q". La concezione classica di un'onda comporta però che nessuna onda W possa avere una precisa posizione spaziale in un tempo specificato. Idealmente, un disturbo d'onda si diffonde nell'intero mezzo in cui si verifica; ma in un punto geometrico un disturbo d'onda non è affatto un disturbo d'onda, bensí semplicemente un impulso periodico. La situazione è simile nel caso del "pacchetto d'onde": la "propagazione" di tale pacchetto è una parte essenziale del fatto che esso possiede proprietà ondulatorie riconoscibili. Perciò, dall'affermazione che W è un'onda (o un "pacchetto" d'onde) possiamo concludere che *W non ha una posizione spaziale determinabile con precisione...* Designiamo questa seconda proposizione con "non-Q". In base a quanto abbiamo detto finora, non-Q non è la negazione di Q; Q si riferisce a P, mentre non-Q si riferisce a W. Quando però P e W sono usate congiuntamente per indicare la medesima entità, avremo la notazione Q e non-Q. Il nostro professore [cfr. § 1.1] deve rappresentare ogni sfera data come una PW, dal momento che essa possiede proprietà tanto corpuscolari quanto ondulatorie: è questa la situazione concettuale che le sue osservazioni gli hanno imposto» [Hanson 1958, trad. it. p. 167]. Ma questa situazione, «logicamente intollerabile» [*ibid.*, p. 168; cfr. del resto l'espressione di Neumann «la autocontraddittoria natura duale della luce»] per il professore ottocentesco del § 1.1, sta alla base di una profonda ristrutturazione concettuale (e logica). Non è possibile in questa sede delinearne i tratti; basterà osservare, ancora con Hanson, che «la massa puntiforme, che è una concezione primariamente cinematica, è il punto di partenza della teoria classica dei punti materiali. L'impulso d'onda, una concezione primariamente dinamica, è il punto di partenza della teoria quantistica. I linguaggi derivanti da una tale diversa origine dovrebbero rivelare da cima a fondo questa differenza logica: e cosí in effetti è» [*ibid.*, p. 181]. Il che comporta anche un riorientamento disciplinare: «Nella fisica classica, gli studi di cinematica sono anteriori a quelli di dinamica; nella fisica quantistica questa divisione e quest'ordine non sono realizzabili. Le qualità primarie erano fondamentali per le concezioni statico-cinematiche che la meccanica classica costruí all'interno di uno spazio euclideo; le proprietà dinamiche dei corpi erano ancillari alle prime. Ora le proprietà statistico-cinematiche di una particella subatomica sono determinate dalle sue proprietà dinamiche; la dinamica quantistica è la disciplina che gode di una posizione prioritaria. Il concetto fondamentale della microfisica è l'*interazione*. L'atomo democriteo-newtoniano-daltoniano non può spiegare le osservazioni realizzate nel nostro secolo. Esso postulava proprietà – come l'impenetrabilità, l'omogeneità, la sfericità – che non ordinano e non integrano piú i nostri dati; per rendere ragione di tutti i fatti l'atomo dev'essere un sistema complesso di enti piú fondamentali. Sono stati scoperti elettroni, protoni, neutroni, positroni, mesoni, antiprotoni, antineutroni e fotoni dei raggi γ; altre particelle saranno probabilmente scoperte se si vorranno spiegare certe "lacune" nei nostri esperimenti. Ma le nuove particelle non possono essere le particelle puntiformi della filosofia naturale classica» [*ibid.*, p. 146].

1.3. Visibile e non visibile. Osservativo e teorico.

Ma dell'antico atomismo democriteo un aspetto non indifferente è pervenuto fino alle attuali concezioni delle particelle. Il nome di Democrito resta infatti intimamente legato all'idea secondo cui l'essere si trova disseminato in una moltitudine quasi infinita di frammenti semplicissimi, ben localizzati nello spazio, raggruppati in un piccolo numero di tipi e in grado di esercitare gli uni sugli altri delle interazioni usualmente localizzate, o di contatto o almeno decrescenti con l'aumentare della distanza. Si potrebbe anche aggiungere che nulla sembra piú semplice da visualizzare di un «piccolo granello» come l'atomo democriteo. Eppure proprio una locuzione come «piccolo granello» è spia di un'autentica difficoltà: è tale piccolezza, infatti, che sembra precludere agli «atomi» la possibilità di una osservazione diretta. Essi sono dunque *invisibili*.

Ma a questa invisibilità si potrebbe arrivare per un'altra via, meno diretta, ma con esiti piú radicali. Gli «atomisti», usando il termine in un'accezione filosoficamente ampia, hanno nutrito opinioni molto diverse sul problema di quelle che dovevano essere le proprietà piú di rilievo degli atomi, oltre (ovviamente) la loro indivisibilità. La posizione aveva importanza centrale per Democrito, la figura per Epicuro e Lucrezio, Gassendi s'interessò soprattutto alle proprietà combinatorie, Newton ai moti degli atomi, Boyle fu attratto dall'insolubilità, mentre la massa colpí soprattutto l'attenzione di Lavoisier (conservazione della massa), Richter (proporzioni costanti) e Dalton; Berzelius (teoria della valenza) notò soprattutto la forza di unione degli atomi e Proust (insistendo ancora sulla indivisibilità) fece degli atomi d'idrogeno le pietre da costruzione ultime della materia. Faraday e Weber, Mayer e Clausius, Boltzmann e Maxwell, ecc. sottolinearono a loro volta altre proprietà, «ma tutti gli atomisti si servirono delle pro-

prietà da loro prescelte nello stesso modo, ossia per spiegare le proprietà macrofisiche dei corpi riconducendole ai caratteri di queste particelle componenti» [*ibid.*, p. 235]. È proprio questo che, *in linea di principio*, sembra rendere gli atomi non raffigurabili. Se gli atomi democritei, infatti, sono *per ipotesi* il risultato della divisione fino all'ultimo stadio possibile, essi non vanno soggetti in se stessi a nessuno dei mutamenti sperimentati dalla percezione sensoriale (è questa l'interpretazione data da Hanson): ciò a priori garantisce che essi *non* possono venir osservati. Come scrive Copernico nel *De revolutionibus orbium coelestium* (libro I, cap. VI), si tratta di corpuscoli «minimi e indivisibili» che, non essendo «sensibili», sfuggiranno quindi sempre all'osservazione. Ovvero, per dirla con Francesco Bacone e poi con Newton, essendo il colore una proprietà «secondaria», gli atomi sono «trasparenti» quindi *invisibili*, ecc. Perché non tendere allora questa «trasparenza» fino a farne una sorta di *trasparenza di puri concetti ausiliari*? Cosí fecero Ostwald, Mach e Karl Pearson e altri dopo di loro. Ostwald sosteneva che gli atomi andavano addirittura banditi dalla scienza fisica, Mach vi vedeva solo delle analogie, Pearson li considerava «*concezioni intellettuali* con l'aiuto delle quali i fisici possono classificare i fenomeni» [1911, p. 85]. Vaihinger, piú radicale ancora, affermava: «Si è visto di recente che l'atomo è una finzione, una connotazione finzionale» [1911, trad. it. p. 78].

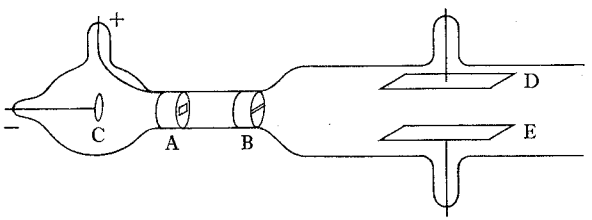


Figura 3.

Mentre ancora nel 1897 William Thomson (Lord Kelvin) sosteneva che l'elettricità è un liquido continuo omogeneo anziché avere una struttura granulare, contemporaneamente Joseph John Thomson riusciva a eseguire il suo celebre esperimento in cui determinò la carica specifica e/m per i raggi catodici. Nella figura è schematizzato il dispositivo di Thomson. I raggi catodici provenienti dal catodo C vengono fatti passare attraverso le fenditure A e B, che servono a restringere il fascio. Il fascio passa quindi attraverso la regione tra D ed E e viene osservato sullo schermo graduato all'estremità destra del tubo. Caricando le armature D e E si produce una deflessione del fascio verso l'alto o verso il basso. Il verso della deflessione mostra che la carica portata dal fascio è negativa. Generato opportunamente un campo magnetico nella regione tra D ed E, di direzione perpendicolare al piano del foglio, si osserva che anche tale campo produce una deflessione del fascio verso l'alto o verso il basso, in accordo con la carica negativa dei raggi. Compensando le deflessioni prodotte dai campi elettrico e magnetico, risulta possibile calcolare la velocità del fascio. L'ampiezza della deflessione prodotta da ciascun campo separatamente dà una misura di e/m, rapporto tra carica e massa dei costituenti del fascio. Il valore ottenuto da Thomson per i raggi catodici era molto piú elevato, per un fattore di diverse migliaia, dei corrispondenti valori determinati per gli ioni dell'elettrolisi. Thomson ne concluse che il raggio catodico consisteva di particelle con massa molto piú piccola di quella degli ioni e con carica negativa. Li battezzò «corpuscoli» e chiamò «elettrone» la loro carica. In seguito vennero chiamate *elettroni* le particelle stesse.

Si riconsiderino però le prime esperienze miranti ad avvalorare la natura granulare della materia (impiego dei raggi X per studiare la natura dei cristalli proposto alla fine del secolo scorso da Max von Laue) della carica elettrica (Millikan, 1910) o lo stesso procedimento impiegato nel 1899 da Thomson per la misurazione della carica specifica, cioè del rapporto tra carica elettrica e massa di singole particelle (nella fattispecie di quei «granuli di elettricità» che erano gli elettroni e gli ioni (per il dispositivo di Thomson, cfr. fig. 3)). Non solo queste procedure sperimentali hanno permesso di articolare ulteriormente concezioni della materia e dell'elettricità emerse nell'Ottocento (per questi aspetti si vedano gli articoli «Atomo e molecola» e «Materia» in questa stessa *Enciclopedia*), ma, poiché i valori numerici di massa, energia, carica elettrica via via determinati non sono il risultato di una interpretazione di quantità mediate su grandi numeri, ma valori consistentemente trovati su ogni singola particella, esse stanno all'origine di una vera e propria fisica delle particelle «elementari».

Il passo successivo è stato rappresentato dalla «rivelazione» delle particelle. Le scoperte che fecero seguito a quelle della radioattività naturale (Becquerel, 1896; i Curie, 1898) e della radioattività indotta (Rutherford, 1919) furono il risultato di un simultaneo progresso teorico e strumentale. Ora, il passo fondamentale nello sviluppo degli strumenti è quello dovuto a Wilson, che è riuscito a mostrare (1912) la possibilità di rivelare il percorso di singole particelle cariche in gas contenenti vapori saturi con la cosiddetta camera a nebbia (o camera di Wilson). Questo dispositivo è costituito da un recipiente contenente appunto un gas e un vapore in condizioni termodinamiche assai vicine alla sua condensazione. Se non sono presenti «nuclei di condensazione» la condensazione non avviene; ma se una particella carica la attraversa, producendo una scia di ioni positivi ed elettroni, questi possono agire come nuclei di condensazione. Abbassando rapidamente la temperatura mediante una espansione adiabatica, la «traccia della particella» viene segnata da una successione di goccioline tanto piú densa quanto maggiore è il potere ionizzante della particella.

A questa prima «finestra sul mondo fisico» (come fu detta) ben presto altri strumenti di rivelazione si sono affiancati (per esempio, emulsioni fotografiche, introdotte da Marietta Blau e Hertha Wanbacher nel 1937, ma sostanzialmente perfezionate solo nel dopoguerra soprattutto per merito di Occhialini e Powell). Ma bastino tali sommari cenni; preme qui sottolineare che tali particelle, pur direttamente non osservabili, sono dunque «rivelabili»: la loro esistenza (come nota Hanson [1958, cap. VI]) può venir inferita retroduttivamente dalle loro «tracce». Si tratta quindi di una «osservazione» indiretta su cui si può innestare anche una concezione realistica delle particelle. Scrive per esempio Shapere: «Dall'opera teorica di Ampère e di Weber fino a quella di Lorentz, dalla ricerca sperimentale di Faraday sull'elettrolisi all'esperimento di Millikan, c'è stato un accumulo di ragioni per sostenere che l'elettricità viaggia in unità discrete... È un'ironia della sorte che il termine 'elettrone' – spesso preso nei manuali della tradizione positivista come un esempio di termine teorico – abbia proprio uno stato del genere» [Shapere 1969, p. 156]. Forse, «concezioni intellettuali» o «connotazioni finzionali» sono un po' troppo «trasparenti» o «rarefatte», se può

essere consentita la metafora, per lasciare tracce in un rivelatore, ecc. [cfr. *ibid.*]. Una versione della dicotomia teorico/osservativa piú sofisticata di quella di Mach, Pearson, Vaihinger, ecc. e forse anche di quella elaborata dall'empirismo logico, è probabilmente necessaria [cfr. ancora, in questa direzione, le tesi avanzate da Shapere 1969]. Ma va aggiunto subito che, in questo paragrafo, la discussione circa l'entità «elettrone» ed eventuali altri *inosservabili* (ma «osservabili» nel senso indiretto e sofisticato qui introdotto) si è svolta ancora nel quadro concettuale classico. Nel contesto quantistico l'interazione osservazione-teoria è piú complessa. Si faccia ancora il caso specifico dell'elettrone. Nota a questo proposito Hanson [1958]: «Nell'immediata prossimità di un elettrone il gradiente di potenziale è un miliardo di volte maggiore di qualsiasi cosa noi possiamo esaminare sperimentalmente. Born esplorò la possibilità che in coincidenza con un tale enorme gradiente di potenziale le leggi dell'elettrodinamica richiedano di fatto una modificazione (1934); queste leggi modificate dovrebbero conservare la validità a qualunque distanza per quanto piccola dal centro dell'elettrone. Ma qui non c'è in realtà un elettrone, non nel senso di una particella infinitesimale dotata di massa: è piuttosto qualcosa di simile a un punto matematico al centro di un campo elettrico di un'intensità incredibile. È questa la ben nota *singolarità*. Born dimostra che è possibile soltanto un tipo di singolarità (almeno nella sua teoria); perciò tutti i punti singolari che ci si presentano devono essere completamente uguali. Una qualche cosa che non potesse essere descritta da qualcosa di simile alle equazioni non lineari di Born semplicemente non sarebbe un elettrone; ma ogni particella che *sia* descrivibile in questo modo dev'essere identica con ogni altra particella che possa essere descritta in questo modo. L'osservazione di elettroni è chiaramente un'attività sofisticata, sostenuta dalla teoria. È l'osservazione di enti che, in linea di principio, sono particelle-onde non visualizzabili e assolutamente identiche» (trad. it. p. 160).

2. *Le basi.*

Si assumerà che il lettore sia familiare con i primi sviluppi della teoria quantistica, dai quanti planckiani e dai fotoni di Einstein fino alla meccanica delle matrici di Heisenberg, Born, Jordan e alla meccanica ondulatoria, e con la modellistica atomica (modelli classici, atomo di Bohr, ecc.): egli troverà comunque una trattazione di questi argomenti in altri articoli di questa stessa *Enciclopedia* (cfr. in particolare l'articolo «Materia», vol. VIII). In questo paragrafo si delineeranno quelle scoperte che hanno permesso il costituirsi di una vera e propria fisica delle *particelle* come disciplina dotata di una sua consistenza e autonomia. Questo consolidarsi dello studio delle particelle «elementari» è avvenuto, si badi, quando «l'età eroica» della fisica atomica si era ormai conclusa con le grandi acquisizioni del 1924-27, quel «periodo creativo» la cui «ricostruzione storica» richiederebbe, forse, «la stessa altissima perizia della storia di Edipo o di Cromwell», come dice Oppenheimer [1954, trad. it. p. 142].

2.1. La scoperta del positrone e l'antimateria.

Verso il 1930 la struttura dell'atomo pareva dunque compresa: si sapeva che vi era un nucleo, a carica positiva, situato al centro e che era circondato da elettroni, cioè da cariche elettriche «elementari» negative. Una volta stabilita l'equazione di Schrödinger, l'indagine di tale struttura avrebbe dovuto dunque diventare un puro problema matematico, trattandosi sostanzialmente di trovare soluzioni dell'equazione di Schrödinger rappresentanti gli orbitali dei diversi elettroni... È però ben noto che in questo modo si riesce a spiegare perfettamente solo la struttura dell'atomo piú semplice, quello di idrogeno, mentre per atomi piú complessi la sola equazione di Schrödinger non è sufficiente. Per risolvere alcune difficoltà (per esempio: perché tutti gli elettroni non vanno ad occupare l'orbitale piú vicino al nucleo?) occorrono alcune ipotesi ad hoc.

Inoltre l'equazione di Schrödinger *non è relativistica*, non è cioè invariante per trasformazioni di Lorentz. Si deve a Dirac (1928) l'aver scritto per gli elettroni una generalizzazione dell'equazione di Schrödinger che rende conto dei dati sperimentali e fa sparire qualche ipotesi ad hoc: ne segue infatti che l'elettrone deve possedere uno spin o momento angolare intrinseco (uguale a $h/4\pi$) e un momento dipolare magnetico, ecc. Ma, soprattutto, dall'equazione di Dirac scaturisce una nuova direzione di ricerca.

Dirac stesso notò infatti che la sua equazione dell'elettrone era dotata di due tipi possibili di soluzione, il che sembrava rappresentare per l'elettrone un'energia rispettivamente positiva o negativa. Ora, sembrava impossibile interpretare le soluzioni indicanti una energia cinetica *negativa* se non in termini di una particella identica all'elettrone nella massa, ma dotata di carica elettrica opposta e di proprietà diverse. Dirac suppose, piú precisamente, che nell'universo osservato ordinariamente gli stati a energia negativa degli elettroni fossero occupati con un elettrone in ciascuno di essi. Per il principio di esclusione (Pauli, 1925: in breve, data una configurazione di ψ soluzione dell'equazione di Schrödinger, in essa non possono trovarsi piú di *due* elettroni), nessun salto quantico è ammesso tra questi stati. Esiste quindi un mare di elettroni con $E<0$, ma di essi normalmente non ci si accorge, in quanto l'unico modo di rivelarne la presenza sarebbe un salto quantico. Dirac osservò allora che, se anche è elevatissima l'energia da comunicare a un elettrone per farlo «saltare» da uno stato di energia negativa a uno di energia positiva, non si potevano escludere esperienze ad hoc in cui si fornisce a un elettrone con $E<0$ l'energia sufficiente per passare a uno stato con $E>0$. In casi del genere rimaneva un «buco» nel mare di elettroni a energia negativa, una *mancanza* interpretabile come una *presenza* di un «elettrone di carica positiva».

Si trattava, almeno fino a questo punto, di una previsione teorica, cioè di «poco piú che una speculazione» come lo stesso Dirac ha osservato retrospettivamente [1974, p. 312]. Ma è fin troppo noto (cfr. del resto il citato articolo «Materia», alle pp. 911 sgg.) che l'elettrone positivo o positrone è stato individuato per la prima volta (1932) da Carl D. Anderson (con la camera di Wilson) nella

radiazione cosmica. È forse meno noto che Anderson era completamente all'oscuro della predizione tratta dalla teoria di Dirac: la storia di come una anticipazione teorica e una «osservazione» sperimentale vengono a convergere determinando una nuova «entità» è stata mirabilmente ricostruita da Hanson [1963].

Qui due aspetti preme rilevare. Il primo concerne la questione dell'«antimateria». Quando un elettrone e un positrone s'incontrano, si annichilano. L'elettrone cade nel buco di energia negativa e scompare. L'energia corrispondente che, come già si è detto, vale almeno $2mc^2$ viene emessa sotto forma di fotoni (raggi γ). Inversamente, un fotone sufficientemente energetico può in opportune condizioni sparire dando luogo a una coppia elettrone piú positrone. Senza seguire una stretta esposizione cronologica, ma anticipando, va detto che quello dell'elettrone non è un caso eccezionale: ogni particella ha la sua antiparticella, che sta ad essa come il positrone sta all'elettrone. L'equazione di Dirac regola soltanto il comportamento delle particelle con spin 1/2. Ma analoghe equazioni relativistiche sono state scritte per particelle di spin diverso. Cosí, per esempio, esiste l'equazione di Klein-Gordon per le particelle di spin zero e l'equazione di Proca-Yukawa per quelle di spin 1. Tutte queste equazioni, per il caso non relativistico di bassa velocità, equivalgono sostanzialmente all'equazione di Schrödinger, con l'aggiunta delle ipotesi ad hoc dello spin. Ma non ci si addentrerà qui in questi aspetti della teoria. Si concluderanno questi brevi cenni circa l'antimateria riprendendo un'osservazione di Emilio Segrè, lo scopritore dell'antiprotone (1955): «La simmetria tra particelle e antiparticelle è una delle nuove verità della fisica. La relazione tra particelle e antiparticelle richiede che esse abbiano esattamente la stessa massa, la stessa vita media quando decadono liberamente, lo stesso momento angolare, carica elettrica eguale e opposta e momento magnetico eguale e orientato oppositamente rispetto al momento angolare, stranezza eguale e opposta. Anche altre proprietà su cui non indugiamo sono o eguali o opposte. Se sostituiamo in un nucleo ogni protone con un antiprotone e ogni neutrone con un antineutrone e poi gli mettiamo attorno positroni, abbiamo formato un atomo di antimateria. Con gli antiatomi si possono poi formare addirittura antimondi. Per un abitante di un antimondo tutto sarebbe eguale all'osservato di un abitante di un mondo. Non vi sono differenze intrinseche e l'osservazione astronomica è incapace di distinguere un mondo da un antimondo. Se però materia e antimateria si incontrano, si annichilano ed entro un brevissimo tempo tutta la loro energia è trasformata in neutrini, antineutrini e raggi gamma che si allontanano dal luogo del disastro con la velocità della luce» [1976, pp. 247-48].

Il secondo riguarda lo stesso attributo «elementare». Heisenberg ha osservato che, con Dirac, «è la nostra concezione globale delle particelle che si trova modificata». Tale scoperta va dunque nella direzione di un'ulteriore varianza del significato di particella. Aggiunge Heisenberg: «Fino a quel momento ogni fisico ha immaginato le particelle elementari in accordo con la filosofia di Democrito, considerandole cioè come entità non modificabili, che esistono realmente in natura e che restano sempre identiche, senza possibilità di cambiamento e di trasformazione in qualcosa d'altro... Ma in seguito alla scoperta di

Dirac tutto appare differente, in quanto ci si può chiedere perché un protone non potrebbe essere talvolta un protone piú una coppia comprendente un elettrone e un positrone, e cosí via... È per questo che il problema della divisibilità della materia ci appare sotto luce nuova. Finora si pensava che non ci fossero che due alternative: o dividere la materia illimitatamente ottenendone delle porzioni sempre piú piccole oppure non ammettere tale suddivisione all'infinito, ottenendo invece le particelle ultime. Adesso scorgiamo una terza possibilità: possiamo dividere la materia mediante un procedimento ripetitivo ma non pervenire lo stesso alle particelle minime, in quanto creiamo delle particelle trasformando dell'energia» [citato e discusso in Amaldi 1979, pp. 54-55].

2.2. Particelle reali e virtuali.

L'assoggettamento dello stesso campo elettromagnetico alle regole della meccanica quantistica (inizialmente Born, Heisenberg, Jordan, Dirac, 1927; quindi Bohr, Rosenfeld, 1933) ha condotto, sul finire degli anni '20 e nei primi anni '30, a una quantizzazione del campo elettromagnetico in cui sono emerse limitazioni analoghe a quelle del principio di Heisenberg.

In breve, la teoria quantistica della radiazione elettromagnetica può venir cosí delineata. Si considerino, dapprima, un sistema di atomi e un campo elettromagnetico, astraendo dalla loro interazione. Ciascun atomo sarà in un determinato livello energetico e altrettanto avverrà per ciascun grado di libertà del campo. Questa situazione può prolungarsi indefinitamente e costituisce uno stato stazionario del sistema. Ma il campo elettromagnetico e gli elettroni interagiscono e questa interazione può essere trattata come una perturbazione dello stato stazionario. Essa ha per effetto di far passare ogni tanto un atomo da un livello energetico E_m a un altro E_n, facendo contemporaneamente passare un grado di libertà del campo da un livello energetico a uno *contiguo*, con conservazione dell'energia totale. Si avrà dunque $E_m - E_n = h\nu$, il fotone venendo emesso se $E_m > E_n$ e assorbito se $E_m < E_n$.

Ora: la meccanica ondulatoria era nata per una esigenza di simmetria fra il campo elettromagnetico, del quale si era scoperto un aspetto corpuscolare, e le ordinarie particelle, delle quali, per converso, veniva fatto di sospettare un aspetto ondulatorio. La quantizzazione del campo elettromagnetico sembra rompere questa simmetria. Se i fotoni sono i quanti del campo elettromagnetico, non dovrebbero gli elettroni a loro volta essere i quanti del loro campo ondulatorio ψ? Ebbene, Jordan e Wigner riuscirono a dimostrare che la ψ che entra nelle equazioni di Dirac può essere quantizzata in modo analogo al campo elettromagnetico. È la cosiddetta seconda quantizzazione. In questo modo gli elettroni e i positroni divengono in un certo senso i quanti del campo ψ e, come succede per i fotoni, possono essere emessi o assorbiti (ma soltanto a coppie) nei processi d'interazione.

Il corpo di dottrina che ha per oggetto la teoria del campo elettromagnetico e del campo elettronico quantizzati e interagenti, ha ricevuto il nome di elettrodinamica quantistica. Essa, attraverso i perfezionamenti apportati da Fermi, Pauli,

Tomonaga, Schwinger, Feynman, Dyson e molti altri è diventata uno strumento fra i piú precisi della fisica moderna. Il suo successo ha comportato, com'è noto, il riorientamento finale circa la dicotomia di «particella» e «campo». Intesi non piú come entità distinte, particella e campo diventano ora le facce di una stessa medaglia, di una sorta di «sostanza» fisica che può propagarsi sotto forma di onda attraverso lo spazio portando energia, quantità di moto, massa, carica, ecc. Una particolarità di ogni *campo* è il fatto che a esso è associata una *massa* ben precisa. Cosí un grumo del campo elettronico può venir creato con una energia e una quantità di moto qualsiasi, ma possiede sempre la stessa massa invariabile, che è proprio la massa della particella che si suole chiamare elettrone! L'interazione di un campo con un altro accentua le proprietà corpuscolari dei campi: è dunque possibile parlare di una particella come «creata» in un dato punto, «in moto» verso qualche altro punto e quindi annichilata. Contemporaneamente è possibile anche una rappresentazione piú complessa: un grumo di energia viene creato, si propaga come un'onda verso un punto e quindi viene di nuovo completamente assorbito.

Si consideri ora un aspetto piú specifico: dalla relazione di Heisenberg, scritta nella forma $\Delta E \Delta t \geq \hbar$ deriva che per controllare con precisione la conservazione dell'energia, rispettata nella meccanica quantistica, occorre disporre di *tempi lunghi*. L'energia di una particella è dunque indeterminata e la conservazione dell'energia vale solo su tempi lunghi. Ma energia ormai significa massa (per l'equivalenza relativista $E = mc^2$): la massa di una particella può dunque «fluttuare» in modo tanto piú cospicuo quanto piú piccolo è l'intervallo Δt che

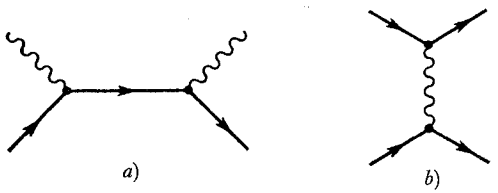


Figura 4.

Diagrammi di Feynman. Le linee piú spesse corrispondono agli elettroni, quelle ondulate ai fotoni. Le linee elettroniche sono provviste di una freccia; le linee ondulate, invece, non hanno freccia in quanto i fotoni sono le antiparticelle di se stessi. Il diagramma *a*) rappresenta l'urto di un elettrone con un fotone e la conseguente diffusione di ambedue (effetto Compton). «Come si vede il processo s'interpreta cosí: l'elettrone incidente assorbe il fotone incidente (vertice sinistro) e si converte in un elettrone virtuale (linea interna), questo emette un altro fotone (vertice destro), ritorna un elettrone reale e si allontana. Lo sperimentatore non si accorge che i prodotti finali sono *un altro* elettrone e *un altro* fotone rispetto a quelli incidenti... È ovvio poi che lo stesso diagramma (letto da destra a sinistra) può rappresentare l'urto di un positrone e di un fotone. Ma può anche rappresentare l'annichilazione di una coppia, con la creazione di due fotoni. Un elettrone viene da sinistra e un positrone da destra; l'elettrone nel vertice di sinistra emette un fotone e si muta in un elettrone virtuale, il quale a sua volta si annichila nel vertice di destra col positrone, emettendo un secondo fotone» [Toraldo di Francia 1976, p. 436]. Il diagramma *b*) può rappresentare l'urto di due elettroni (o di due positroni). Il primo elettrone emette un fotone virtuale che viene assorbito dal secondo elettrone.

si considera. Dunque in un sistema possono venire emesse o generate particelle anche quando il sistema *non* possiede l'energia necessaria, purché tali particelle sussistano per intervalli molto brevi e vengano subito riassorbite dal sistema. Si parla allora di particelle virtuali: se ne darà un cospicuo esempio nel § 3.1 (mesone di Yukawa).

Si accennerà infine, brevemente, alla rappresentazione grafica, suggerita da Feynman nel 1949 per i calcoli perturbativi e oggi universalmente accettata. Un diagramma o grafico di Feynman consiste di linee esterne, di linee interne e di vertici. I vertici sono i punti in cui le linee si incontrano; le linee esterne, ciascuna delle quali ha un solo vertice in comune con il resto del diagramma, rappresentano le particelle iniziali (o incidenti) e le particelle finali (o generate, o diffuse); le linee interne congiungono ognuna due vertici e rappresentano le particelle virtuali; i vertici, infine, rappresentano i processi elementari di interazione.

Un esempio è dato dai diagrammi riportati nella figura 4. Ma va osservato che a uno stesso fenomeno fisico possono corrispondere *infiniti* diagrammi di Feynman, sempre piú complicati. Come questo sia possibile può essere chiarito dalla figura 5 e dal confronto con la figura 4.

«Non possiamo abbandonare l'argomento senza fare qualche considerazione di carattere generale. Lo sbalorditivo successo dell'elettrodinamica quantistica è noto da molto tempo. Eppure coloro che commentano o criticano i principî della meccanica quantistica raramente si riferiscono ad essa, rimanendo piuttosto ancorati alle perplessità e alle battaglie degli anni venti o trenta. È mai possibile continuare ad additare come non *digeribile* il fatto che una particella non abbia contemporaneamente posizione e quantità di moto ben definite, quando dobbiamo accettare fatti tanto piú lontani dal pensiero ingenuo, come quello che ogni regione dello spazio pullula di (infinite!) particelle fantasma, che appaiono e scompaiono senza essere in alcun modo localizzabili?» [Toraldo di Francia 1976, p. 439].

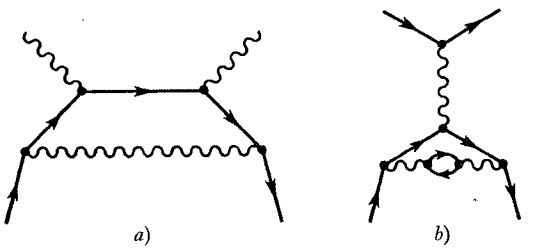


Figura 5.

Diagrammi di Feynman. In *a*) (rispettivamente in *b*) un diagramma piú complicato che in *a*) (rispettivamente in *b*) nella figura 4 per l'urto di un elettrone con un fotone (rispettivamente l'urto di due elettroni o di due positroni). In *a*) l'elettrone incidente, prima di emettere il fotone reale, emette un fotone virtuale, che poi viene assorbito dall'elettrone diffuso. In *b*) uno degli elettroni incidenti emette preliminarmente un fotone virtuale, il quale dà luogo a una coppia virtuale che si annichila creando di nuovo un fotone virtuale. Quest'ultimo viene infine assorbito da uno degli elettroni finali. (Da Toraldo di Francia 1976).

2.3. Il nucleo atomico.

È noto che il nucleo di alcuni elementi può emettere particelle α (nuclei di elio), particelle β (elettroni) e fotoni γ. Una volta chiarita la struttura dell'atomo (cfr. l'articolo «Atomo e molecola» in questa stessa *Enciclopedia*) in termini di nucleo ed elettroni, era difficilmente concepibile una particella «elementare» che potesse avere queste proprietà. Il nucleo è dunque una struttura composta. Ma come? L'idea piú semplice è la seguente: un nucleo di dato numero atomico Z ha carica positiva uguale in valore assoluto a Z volte quello dell'elettrone (cfr. il § 6 dell'articolo «Materia» in questa stessa *Enciclopedia*); il nucleo dell'idrogeno, cioè il protone, ha una carica positiva uguale (in valore assoluto) a quella dell'elettrone; i nuclei degli elementi successivi sono formati da un numero A di protoni, piú A − Z elettroni. Si ottiene cosí una carica complessiva di $Ae-(A-Z)e=Ze$ e un peso approssimativo uguale ad A volte quello di un protone (data la piccolezza della massa dell'elettrone).

Si consideri ora il principio di indeterminazione, questa volta nella formulazione $\Delta x \Delta p_x \geq \hbar$. Se un elettrone è confinato nel nucleo che risulta avere dimensioni dell'ordine di 10^{-13} cm, l'indeterminazione della sua posizione non supera questo valore; dunque Δp_x non può essere inferiore a $\hbar/10^{-13}$, deve dunque avere almeno l'ordine di grandezza 10^{-14} nel sistema CGS. La quantità di moto p dell'elettrone non sarà dunque inferiore a 10^{-14}, in quanto una grandezza non può essere minore della sua indeterminazione. L'energia dell'elettrone è del resto data da $E=c\sqrt{p^2+m^2c^2}$ e poiché la massa m dell'elettrone è circa 10^{-27} g, si ottiene per l'energia un valore non inferiore a circa 10^{-4} erg o a 100 MeV (1 MeV è 10^6 eV, e 1 elettronvolt, eV, è l'energia che acquista un elettrone passando attraverso la differenza di potenziale di 1 volt; essa vale $1{,}6 \times 10^{-12}$ erg). Una stima del genere non permette di comprendere perché l'elettrone debba restare nel nucleo; del resto gli elettroni nell'emissione β escono con energie molto minori, dell'ordine del MeV.

Questa e altre difficoltà furono rimosse dalla scoperta del neutrone, scoperta che «ha una storia complicata, drammatica e interessante» [Segrè 1976, p. 175]. Basterà qui dire che il successo di Chadwick nel 1932 doveva ben presto richiamare l'interesse su questa sorta di «protone neutro», o neutrone come venne in seguito abitualmente chiamato, dotato della stessa massa dell'abituale protone e di spin 1/2. In altri articoli della presente *Enciclopedia* (cfr. per esempio «Atomo e molecola», «Materia», ecc.) si è insistito sull'importanza degli isotopi, cioè degli atomi della stessa specie chimica di peso diverso, la cui esistenza viene appieno spiegata dalla scoperta di Chadwick. La specie chimica è data dal numero Z dei protoni (uguale al numero degli elettroni periferici): ma il nucleo è composto da Z protoni e N neutroni (cioè, come si dirà, da $A=Z+N$ nucleoni), pertanto ha una carica (positiva) Ze e un numero di massa A, cioè un peso pari ad A volte circa il peso dell'atomo di idrogeno.

Qui non s'insisterà, per ovvie ragioni, sulla rilevanza degli isotopi (il deuterio o idrogeno pesante è individuato, nello stesso anno, da Urey) e sull'ampia ricerca che ne è derivata. Varrà invece la pena di anticipare un punto che verrà ulteriormente ripreso.

«Il modello nucleare a neutroni e protoni sembra definitivo. Ne nasce l'importanza dello studio della interazione protone-protone, neutrone-protone e neutrone-neutrone, con la speranza ultima di poter spiegare le proprietà dei nuclei in base a queste interazioni» [Segrè 1976, p. 182]. Tra i successi parziali di questo programma va segnalato il fatto che esperienze sempre piú estese sull'urto nucleone-nucleone hanno messo in luce l'uguaglianza dell'interazione specificatamente nucleare tra protone e protone, neutrone e protone, e neutrone e neutrone. Heisenberg, Condon e Wigner, dagli anni '30 in poi, hanno formulato questo fatto teoricamente considerando neutrone e protone come la stessa particella, ma con valori diversi di una coordinata interna, proprio come un elettrone con lo spin parallelo o antiparallelo a un campo magnetico. Questa coordinata interna, detta spin isotopico, consente un'interessante rappresentazione astratta; matematicamente è assai simile al momento angolare come trattato nella meccanica quantistica. Ha tre componenti e la terza componente T_3 è connessa con la carica elettrica della particella dalla relazione $Q=T_3+B/2+S/2$ ove Q è la carica della particella, B il cosiddetto numero barionico e S la stranezza (cfr. ancora l'articolo «Materia», alle pp. 918-20). Ma su questo si tornerà nei §§ 3.5 e 3.6.

3. *L'epoca della crescita.*

3.1. Nuove particelle.

Il neutrino. Nell'«anno mirabile», 1932, non pochi pensavano che la lista delle «particelle elementari» fosse ormai chiusa: protoni e neutroni formavano i nuclei; aggiungendo gli elettroni si costituivano gli atomi e il quadro si completava aggiungendo i fotoni e gli eventuali gravitoni (ovvero gli ancor oggi ipotetici e misteriosi quanti del campo gravitazionale).

La storia della radioattività β, come fenomeno imbarazzante e autentica sfida per i ricercatori, è un tipico caso esemplare su cui i filosofi e storici della scienza possono utilmente riflettere [si veda, per esempio, la ricostruzione razionale data in Lakatos 1970, trad. it. pp. 245-50]. Il percorso della «scoperta scientifica» è, anche qui, affascinante. Nella radioattività β, com'è noto, un neutrone (protone) del nucleo emette un elettrone (positrone) e si trasforma in un protone (neutrone). Ma un esame dei risultati sperimentali rivelava che nella reazione di decadimento del nucleo era apparentemente violato il principio di conservazione dell'energia (oltre a quelli della quantità di moto e dello spin).

È per questo che tali esperienze «per poco non divennero il maggior esperimento negativo della storia della scienza» [*ibid.*, p. 245]. Se questo non avvenne e tali esperienze ricaddero al rango di semplici rompicapo di routine, fu grazie a una sorta di «particella ladra», come si disse allora [cfr. *ibid.*, p. 247]. Fu Pauli a formulare l'ipotesi che nel decadimento β venisse emessa tale nuova particella, ma fu Fermi, che la battezzò «neutrino», a costruire (1934) una teoria capace

d'inserirsi con un certo successo nella «competizione teorica» [*ibid.*, p. 250] circa l'emissione β. Il neutrino ha massa e carica nulle (come il fotone) e spin 1/2.

Il neutrino ha un'interazione estremamente debole con le altre particelle; di conseguenza le relative reazioni, quando possono avvenire, sono estremamente rare. Cosí un neutrone libero impiega circa un quarto d'ora (vita media) a decadere in un protone e un elettrone emettendo un neutrino. Per le ragioni esposte al § 3.2, oggi il processo si simbolizza $n \rightarrow p + e^- + \bar{\nu}_e$ chiamando la particella emessa $\bar{\nu}_e$ antineutrino.

Per il bilancio energetico, il neutrone ha una massa pari a 939,54 MeV, mentre il protone ha una massa pari a 938,25 MeV, che sommata a quella di 0,51 MeV dell'elettrone dà 938,76 MeV; rimangono dunque a disposizione 0,78 MeV, che si trasformano nell'energia cinetica dell'elettrone e del protone, e nell'energia dell'antineutrino. Ma queste relazioni valgono per il neutrone libero, che si trasforma in un protone libero (piú elettrone e antineutrino). Quando i nucleoni fanno parte di un nucleo, bisogna tener conto dell'energia di legame nucleare e della repulsione elettrostatica fra protoni. Questo fa sí che nella maggior parte dei casi l'energia necessaria per far avvenire la reazione non è disponibile; allora il nucleo è stabile. Dove invece si ha ancora energia sufficiente, si ha radioattività β.

L'esistenza di neutrini è stata per molto tempo accettata solo perché richiesta dai principî di conservazione. Ma la rivelazione diretta della particella si presenta estremamente difficile, a causa della esiguità dell'interazione debole. Solo nel 1956 Cowan e Reines sono riusciti, usando un potente fascio di antineutrini prodotto da un reattore, a rivelare la reazione inversa del decadimento β, $\bar{\nu}_e + p \rightarrow n + e^+$, in cui un antineutrino urta un protone e ne nascono un neutrone e un positrone.

I mesoni di Yukawa e i pioni. Nei primi anni '30 si era dunque a conoscenza di tre tipi fondamentali di «forze» agenti nell'atomo. La forza elettromagnetica interessava le interazioni tra particelle cariche; l'«interazione debole», secondo Fermi, sembrava rendere conto dell'anomalia rappresentata dal decadimento β; l'ultima, l'«interazione forte», era ritenuta la forza di legame che tiene unito il nucleo. La natura di tale forza restava sconosciuta, ma si sapeva che essa era molto piú intensa della forza elettromagnetica o delle interazioni deboli *e aveva un raggio d'azione brevissimo*, pari a quello del nucleo atomico.

Nel 1935 Yukawa, «allo scopo di elaborare una teoria che rendesse conto della natura dei legami nucleari, partí da una analogia. Egli partí dall'osservazione che il campo elettromagnetico, in accordo con la meccanica quantistica, poteva venire interpretato come composto di quanti che agiscono tra le particelle quali portatori di energia e di quantità di moto aventi valori ben determinati» [Yukawa 1974, p. 198]. Questi quanti, cioè i fotoni, interagiscono con le particelle e possono da queste venir emessi o assorbiti: allo stesso modo, i quanti nucleari di Yukawa sono responsabili mediante scambi virtuali dell'interazione nucleonica.

Per compiere questo passo Yukawa modificò radicalmente la comune concezione del protone: invece di considerarlo, come d'abitudine, una sorta di oggetto inerte, egli lo prospettò, per cosí dire, come un centro di attività continua. Esso «si spezza» in un neutrone e in uno di questi «quanti nucleari» carico, oppure in un protone e in uno di questi «quanti nucleari» elettricamente neutro. Per la legge di conservazione dell'energia, la particella ipotizzata da Yukawa appare una particella virtuale (cfr. § 2.2). Dal principio d'indeterminazione $\Delta t \Delta E = \hbar = 7 \times 10^{-22}$ MeV × s e dall'ipotesi che tale particella viaggi alla velocità della luce restando confinata in un raggio di $1{,}5 \times 10^{-13}$ cm (tale è appunto il raggio d'azione del campo nucleare) si ottiene per Δt un valore molto basso, 5×10^{-24}, e di conseguenza un'energia ΔE di circa 140 MeV. Ciò fornisce una stima della massa della particella: se si assume come unità la massa dell'elettrone, si ottiene che i quanti nucleari di Yukawa si collocano *tra* gli elettroni ($m_e = 1$) e i nucleoni ($m_e = 1800$) avendo una massa circa duecento volte quella dell'elettrone. È per questa ragione che Yukawa li battezzò mesoni, congetturando che essi avrebbero avuto vita media di circa 10^{-8} secondi e sarebbero decaduti spontaneamente in elettroni (o positroni) e neutrini.

Come i fotoni, nemmeno i mesoni di Yukawa devono essere sempre virtuali. Nel 1947 Lattes, Occhialini e Powell scoprirono nei raggi cosmici delle particelle, battezzate pioni, di massa pari a 270 m_e, interagenti fortemente con i nucleoni ed elettricamente cariche. Nel 1949 esperienze attuate mediante acceleratori di particelle indicarono l'esistenza del pione neutro (massa quasi uguale a quella del pione carico, vita media inferiore a 10^{-16} secondi). A quindici anni circa dall'ipotesi iniziale, si erano dunque individuati i quanti nucleari di Yukawa.

I muoni. Le vie della scoperta sono tortuose. Un anno dopo la proposta di Yukawa, Anderson e altri avevano individuato particelle «di massa intermedia» (cioè *mesoni*, nel senso in cui il termine era stato impiegato originariamente da Yukawa: circa duecento volte la massa dell'elettrone) nei raggi cosmici. Inoltre questi «mesoni» presentavano il decadimento previsto dalla teoria: essi decadevano spontaneamente in elettroni (o positroni) e neutrini. Ma risultò che essi interagivano con i nucleoni solo debolmente (e per di piú avevano una vita media di 10^{-6}). Sakata e Tanikawa ipotizzarono allora che esistessero due tipi diversi di mesoni: *a*) quelli del tipo individuato da Anderson, non interagenti fortemente con i nucleoni, e *b*) quelli di un secondo tipo, interagenti fortemente con i nucleoni e con vita media in accordo alle previsioni della teoria di Yukawa, che decadevano infine nei mesoni del primo tipo. Già si è visto come questi ultimi siano oggi noti come mesoni π o pioni; i primi, invece, sono oggi conosciuti come *muoni* μ^+ e μ^-. Essi decadono secondo lo schema

$$\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$$
$$\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu$$

Per ragioni di natura sperimentale si arrivò a concludere che l'elettrone e il muone hanno ciascuno un suo neutrino. Si tratta di particelle differenti, indicate usualmente con ν_e e ν_μ rispettivamente. Anche in questo caso, un'anticipazione: il neutrino del muone è stato osservato come differente dal neutrino dell'elettrone nel 1962.

Complessivamente e, μ, ν_e e ν_μ costituiscono la famiglia delle particelle leggere o leptoni. L'elettrone e il muone sono uguali in tutto e per tutto, salvo che per la massa e la vita media. La differenza di vita media non costituisce un grosso rompicapo: di essa si può fornire una ragione abbastanza plausibile. Un enigma rappresenta invece la differenza delle masse rispettive: contro la massa di 0,5 MeV dell'elettrone, il muone ha una massa di 106 MeV. Non si perviene certo a uno scioglimento della questione movendo dal principio che le differenze di massa nascano dalle differenze delle interazioni delle particelle considerate, per cui se le interazioni forti differiscono, le particelle differiscono notevolmente per quanto riguarda la massa; se le interazioni forti sono le stesse, ma sono differenti quelle elettromagnetiche, la differenza nella massa è molto piú piccola; se le particelle sono indistinguibili rispetto alle interazioni forti e a quelle elettromagnetiche, non hanno alcuna differenza di massa. Infatti muone ed elettrone non posseggono alcuna interazione forte!

Tabella 1.
Lo «zoo delle particelle» relativamente stabili. (Da Toraldo di Francia 1976).

Particella	*Antiparticella*	*Carica*	*Massa (MeV)*	*Numero barionico*	*Spin*	*Isospin T_3*	*Stranezza*	*Vita media (s)*
Fotone								
γ	γ	0	0		1			∞
Leptoni								
ν_e	$\bar{\nu}_e$	0	0		1/2			∞
ν_μ	$\bar{\nu}_\mu$	0	0		1/2			∞
e^-	e^+	$-e$	0,511		1/2			∞
μ^-	μ^+	$-e$	106		1/2			$2{,}21 \times 10^{-6}$
Adroni — *Mesoni*								
π^-	π^+	$-e$	140	0	0	-1	0	$2{,}6 \times 10^{-8}$
π^0	π^0	0	135	0	0	0	0	10^{-16}
π^+	π^-	$+e$	140	0	0	1	0	$2{,}6 \times 10^{-8}$
K^+	K^-	$+e$	494	0	0	1/2	$+1$	$1{,}2 \times 10^{-8}$
K^0	$\bar{K}^0$	0	498	0	0	$-1/2$	$+1$	10^{-10}; 6×10^{-8}
Adroni — *Barioni*								
p	$\bar{p}$	$+e$	938	1	1/2	1/2	0	∞
n	$\bar{n}$	0	940	1	1/2	$-1/2$	0	10^3
Λ^0	$\bar{\Lambda}_0$	0	1115	1	1/2	0	-1	$2{,}5 \times 10^{-10}$
Σ^+	$\bar{\Sigma}^-$	$+e$	1190	1	1/2	1	-1	$8{,}1 \times 10^{-10}$
Σ^0	$\bar{\Sigma}^0$	0	1192	1	1/2	0	-1	$\sim 10^{-20}$
Σ^-	$\bar{\Sigma}^+$	$-e$	1196	1	1/2	-1	-1	$1{,}6 \times 10^{-10}$
Ξ^0	$\bar{\Xi}^0$	0	$\sim$1311	1	1/2	1/2	-2	3×10^{-10}
Ξ^-	$\bar{\Xi}^+$	$-e$	1319	1	1/2	$-1/2$	-2	$1{,}7 \times 10^{-10}$
Ω^-	Ω^+	$-e$	1680	1	3/2	0	-3	$1{,}3 \times 10^{-10}$

Il bosone intermedio. Per analogia con le interazioni elettromagnetiche che avvengono per scambio di fotoni e con le interazioni forti che avvengono per scambio di «mesoni» di Yukawa, si è infine supposto che anche le interazioni deboli avvengano mediante lo scambio di una particella carica W, detta bosone intermedio. Ma il bosone intermedio (o *weakon*, come spesso si dice) non è mai stato isolato; ciò, presumibilmente, dipende dal fatto che la massa del bosone dovrebbe essere notevole (qualche decina di GeV: 1 GeV è uguale a 10^9 eV). Si riaccennerà all'argomento nel § 4.3.

Gli iperoni. Si tratta di altre particelle individuate dopo i pioni. Se ne parlerà nel successivo § 3.2 come di «particelle strane». I nucleoni e gli iperoni vengono comunemente detti barioni (o particelle pesanti). Mesoni e barioni rientrano nella famiglia degli adroni, o particelle sottomesse all'interazione forte. La tabella 1 fornisce un quadro delle particelle «relativamente stabili». Il senso di questo termine verrà chiarito nel paragrafo successivo, ove il lettore troverà anche la spiegazione dei numeri quantici come stranezza, isospin, ecc.

Le risonanze. Se gli anni '50 sono stati il periodo delle «particelle strane» (cfr. quanto detto sopra e, piú oltre, il § 3.2), gli anni '60 hanno aperto la porta a una nuova ondata di particelle di vita ultracorta dette comunemente risonanze. Si tratta di particelle (se è ancora lecito usare il termine data la loro brevissima vita!) che interagiscono fortemente e alle quali non sembra vietato da alcuna legge di conservazione (cfr. quanto si dirà nel paragrafo successivo) di seguire una naturale inclinazione a decadere molto rapidamente.

Per fissare le idee, si consideri il seguente esempio. Un protone e un antiprotone si annichilano creando cinque pioni, secondo lo schema $\bar{p}+p \rightarrow \pi^+ + \pi^+ + \pi^- + \pi^- + \pi^0$. L'interpretazione piú soddisfacente delle tracce sulla camera a bolle sembra essere quella di un processo a due stadi. Al primo stadio viene creata una particella detta η_1^0 ($\bar{p}+p \rightarrow \eta_1^0 + \pi^+ + \pi^-$) e quindi η_1^0 decade: $\eta_1^0 \rightarrow \pi^+ + \pi^- + \pi^0$.

Come si vede, l'esistenza di una risonanza viene quindi inferita dallo studio dei prodotti piú duraturi del suo decadimento.

3.2. Sistematica e classificazione.

Il problema. Già quanto si è detto nel § 3.1 basta a mostrare che non si è realizzata quella particolare versione del programma «democriteo» imperniata su *poche* particelle da cui far dipendere la costituzione di tutta la materia. Se esse sono «elementari», ci sono allora troppi elementi del mondo perché l'aggettivo non perda tutta la sua pregnanza! Ma è possibile, tuttavia, una sorta di sistematica e classificazione delle particelle.

Una prima suddivisione è quella in base al tipo di interazione che una particella esibisce in vicinanza di un'altra particella. Com'è noto (cfr. del resto l'ar-

ticolo «Interazione» in questa stessa *Enciclopedia*, vol. VII, p. 886) i tipi constatati sembrano ridursi a quattro: gravitazionale, elettromagnetica, forte e debole. Ma vi è disparità fra le intensità di tali interazioni. Quelle forti superano in intensità le interazioni gravitazionali di 10^{40}; quelle elettromagnetiche sono invece meno intense delle interazioni forti solo per un fattore prossimo a 100; esse superano però le interazioni deboli di un fattore 10^{13} e queste ultime sono circa 10^{25} volte piú intense delle interazioni gravitazionali che, a buon diritto, andrebbero chiamate «debolissime».

Si riconsideri la tabella 1. Mesoni e barioni, cioè gli adroni, subiscono le interazioni forti, i leptoni subiscono le interazioni deboli, i fotoni e tutte le particelle cariche sono sensibili alle interazioni elettromagnetiche; tutte le particelle (a cui andrebbe aggiunto il gravitone) subiscono infine le interazioni gravitazionali. Questo vale, si badi, anche per particelle che si suppongono prive di massa: alla gravitazione, infatti, è sensibile qualunque forma di energia, che si trovi concentrata come massa oppure no (che i fotoni siano influenzati dalla gravità è stato per la prima volta mostrato dalle osservazioni della deviazione da parte del Sole dei raggi luminosi provenienti dalle stelle, osservazione notoriamente interpretata come conferma della relatività generale, ecc.).

C'è quindi una prima regola di classificazione: quanto piú debole è l'interazione, tanto maggiore è il numero di particelle che essa riguarda. Ma c'è una connessione piú interessante: tra intensità delle interazioni e leggi di conservazione. Quanto piú forte è l'interazione, tanto piú questa è vincolata da leggi di conservazione che limitano le possibili trasformazioni fra le particelle (cfr. anche il § 3.3).

Numero elettronico, barionico e muonico. Ci sono, anzitutto, le quattro leggi classiche di conservazione: energia-massa, momento angolare (compreso lo spin), quantità di moto, carica elettrica. Ma queste non sono sufficienti: occorrono nuovi invarianti proprio per spiegare come processi, possibili come quelli usualmente constatati alla luce dei quattro vincoli testé citati, di fatto non hanno luogo.

A differenza delle quattro leggi citate, già note in contesto macroscopico, altre leggi sono state dunque scoperte direttamente attraverso lo studio delle trasformazioni delle particelle. Una trasformazione tipo $p \to e^+ + \gamma$ è ammissibile dal punto di vista della conservazione dell'energia e della carica. Essa non è piú ammissibile – come il lettore potrà controllare – introducendo nuovi invarianti, i «numeri di famiglia».

Si può attribuire infatti a ciascuno dei barioni (cioè al protone e a tutte le particelle piú pesanti) un numero barionico $+1$, e alle loro antiparticelle il numero -1. Questa convenzione non è senza importanza: essa consente tipici processi di produzione di antiprotoni, tipo $p+p \to p+p+p+\bar{p}$ che a prima vista si potrebbe interpretare come la trasformazione di due barioni in quattro barioni; o processi come $p+\bar{p} \to \pi^+ + \pi^- + \pi^0$ ove due barioni sono apparentemente scomparsi. L'invariante è rappresentato dalla somma *algebrica* dei numeri barionici.

La «famiglia dell'elettrone» contiene solo l'elettrone e il suo neutrino, la «famiglia del muone» solo il muone e il suo neutrino. Il numero elettronico di e^- e di ν_e è $+1$, quello di e^+ e ν_e è -1. Per esempio: il decadimento β (cfr. § 3.1) va scritto nella forma $n \to p + e^- + \bar{\nu}_e$ in quanto a sinistra il numero elettronico è 0 e tale deve essere a destra (il numero barionico, invece, è $+1$ a sinistra e $+1$ anche a destra in quanto e^- e $\bar{\nu}_e$ non sono barioni!) Analoga convenzione per il numero muonico: $+1$ per μ^- e ν_μ e -1 per μ^+ e $\bar{\nu}_\mu$).

È stato detto, proprio a proposito di queste e altre leggi di conservazione tipicamente microfisiche che «il quadro risultante del mondo è un caos quasi senza limiti governato soltanto da un gruppo di leggi di vincolo» [Ford 1963, trad. it. p. 221]. Ma tali leggi già spiegano aspetti rilevanti: per esempio la stabilità del protone, cioè l'impossibilità che esso si trasformi in altre particelle di massa minore, è immediata conseguenza della conservazione del numero barionico (queste particelle avrebbero, inevitabilmente, numero barionico zero). Un analogo ragionamento vale per la stabilità dell'elettrone (cfr. del resto l'articolo «Conservazione/invarianza» in questa stessa *Enciclopedia*).

Particelle strane e «stranezza». Le sette leggi di conservazione finora considerate governano però *tutte* le interazioni fra particelle. Si tratta ora di individuare le leggi di conservazione parziali.

Nel 1947 la lista delle particelle conosciute (includendovi uno dei neutrini e tralasciando il gravitone) era di quattordici: i raggi cosmici rappresentavano la fonte principale di particelle di alta energia, l'acceleratore piú potente spingeva i protoni fino a circa 200 MeV e il rivelatore tipico era costituito dalla camera di Wilson. Appena otto anni dopo la lista delle particelle era giunta a una trentina, erano stati costruiti acceleratori potenti (come il bevatrone di Berkeley che accelerava i protoni a 6000 MeV, cioè a 6 GeV), era stata infine inventata la camera a bolle, un nuovo strumento di rivelazione...

Le sedici nuove particelle sono state classificate in quattro famiglie: mesoni K (detti talora kaoni), particelle Λ, particelle Σ e Ξ. Esse vennero riconosciute per la prima volta attraverso alcune tracce, inspiegate, a forma di V nella camera a nebbia di Rochester e Butler (Manchester): la traccia di un pione incidente si arrestava improvvisamente nella camera e, separate da essa di pochi centimetri, vi erano due V, ciascuna col vertice rivolto verso l'estremità terminale della traccia del pione entrante. Una misura della quantità di moto delle tracce visibili permise di congetturare che due particelle neutre erano state create nel punto terminale della traccia del pione per decadere quindi in coppie di particelle di carica opposta, che fornivano appunto le tracce a V.

Oggi le prime particelle a V osservate sono state identificate come mesoni K neutri o come particelle Λ che decadevano secondo gli schemi

$$K^0 \to \pi^+ + \pi^-$$
$$\Lambda^0 \to p + \pi^-$$

Pochi anni dopo vennero identificate anche le particelle K e Λ cariche, cui si aggiunsero quindi le particelle Σ e Ξ (cfr. tab. 1). Tali particelle hanno una vita media notevolmente lunga.

La «stranezza» sembra consistere proprio nel fatto che, mentre la lunga vita di un pione o di un muone appare spiegabile poiché questi decadono in particelle che interagiscono solo debolmente, nei decadimenti di particelle Λ, K, ecc., come si è visto poco piú sopra, si ottengono particelle che interagiscono fortemente. Perché particelle di questo tipo «non approfittano» immediatamente di tali interazioni forti per convertirsi in particelle tipo protoni e pioni?

Come si è visto, ciò che rende praticamente «eterno» il protone è la legge di conservazione dei barioni. Negli anni '50 Gell-Mann e Nishujima hanno, in via analogica, ideato un nuovo vincolo, la legge di conservazione, appunto, della stranezza. A ciascuna delle particelle – come appare nella tabella 1 – è attribuito un numero intero S che può avere valore 0, ± 1, ± 2, ecc. Talvolta, invece della stranezza S, si usa l'ipercarica Y, che è la somma della stranezza S e del numero barionico B.

Si consideri, come esempio, un processo tipo $p+p \rightarrow p+n+\pi^+$ ove si producono dei pioni. La legge di conservazione della stranezza è soddisfatta: si ha per S il valore zero tanto a sinistra che a destra della freccia. Ma nella produzione di particelle strane, queste non possono che andare a coppia («produzione associata» di Pais). Per esempio, nel processo $p+p \rightarrow p+\Lambda^0+K^+$, la stranezza dei due protoni incidenti è zero, mentre i numeri di stranezza -1 e $+1$ della Λ e del mesone K si annullano a vicenda.

Si è detto che si cercano leggi di conservazione che permettano di articolare meglio la classificazione delle particelle: la legge di conservazione della stranezza risponde allo scopo. Per rendersene conto, si consideri ancora il mesone K. Ora, se la conservazione della stranezza fosse una legge assoluta, il mesone K non potrebbe affatto decadere e si unirebbe alle particelle stabili. Ma poiché la conservazione della stranezza è rispettata solo nelle interazioni forti, e non in quelle deboli, al mesone K è vietato unicamente quel tipo di decadimento troppo rapido che caratterizza le interazioni forti. Esso è solo «relativamente» stabile: poiché le interazioni deboli violano la conservazione della stranezza, esse possono infatti provocare il decadimento del mesone K.

Spin isotopico. Si riconsideri la scoperta del neutrone (§ 2.3): nel 1932 elettrone, fotone e protone erano considerate particelle diverse tra loro; la nuova particella, il neutrone, era in modo evidente simile al protone. Neutrone e protone avevano circa la stessa massa, erano entrambi costituenti dei nuclei e si attraevano reciprocamente con una nuova forza molto potente.

Nello stesso anno Heisenberg interpretò teoricamente protone e neutrone come due stati di una singola particella, detta, appunto, nucleone. Lo strumento essenziale di questa rappresentazione fu costituito dall'introduzione di uno spazio astratto **T**, in cui il nucleone è inteso come un vettore che può essere orientato in un verso o nel verso opposto a seconda che si dia l'interpretazione «protone» o quella «neutrone». In questa rappresentazione il nucleone è un «doppietto», una particella che può esistere solo in due «stati».

L'idea era destinata a notevoli estensioni: man mano che venivano trovate nuove particelle, si scopriva che esse si manifestavano in gruppi omogenei, tipo

protone-neutrone. Cosí i tre pioni π^+, π^- e π^0 di cui al § 3.1 formano un tripletto; le particelle Σ anch'esse un tripletto; i mesoni K, come i nucleoni, un doppietto; la particella Λ resta da sola e costituisce dunque un singoletto.

Il passo logicamente successivo è quello di quantizzare lo spin isotopico T, in particolare la sua terza coordinata T_3. Cosí i singoletti hanno spin isotopico 0, i doppietti 1/2, i tripletti 1, ecc. Si tratta allora di enunciare una legge di conservazione anche per lo spin isotopico, formalmente analoga alla legge della conservazione della stranezza e, come questa, formulata per le sole interazioni forti. Intuitivamente questa legge esprime una sorta di indipendenza dalla carica: protone e neutrone sono sullo stesso piano, e cosí i pioni negativi, positivi e neutri interagiscono (fortemente) nella stessa maniera. Oltre che per le interazioni deboli, ciò non vale però anche per la stessa interazione elettromagnetica: un fotone interagisce con le particelle cariche, ma non con quelle neutre.

3.3. Il «fascino discreto» delle simmetrie.

È noto che «individuando le simmetrie di un problema dinamico si possono raggiungere importanti risultati. Per esempio la circolarità di un'orbita elettronica in un atomo di idrogeno è una conseguenza e un indice della simmetria della forza coulombiana esercitata dal nucleo sull'elettrone. La simmetria in questo caso significa che la forza ha la stessa intensità in tutte le direzioni. Princípi di simmetria simili a questo hanno avuto importanza in fisica classica, ma in meccanica quantistica il concetto di simmetria ha acquistato un interesse ancora maggiore sia in profondità che in estensione... Per dare due esempi: la struttura generale della tavola periodica è essenzialmente una diretta e mirabile conseguenza della simmetria dianzi nominata, l'isotropia della forza di Coulomb; l'esistenza delle antiparticelle... fu anticipata nella teoria di Dirac, che è basata sul principio di simmetria relativistica. In tutti e due i casi, come in altri esempi, la natura sembra giovarsi della semplice rappresentazione matematica delle leggi di simmetria» [Chen Ning Yang 1961, trad. it. p. 70].

Uno dei principî di simmetria, la simmetria tra destra e sinistra, è stato discusso fin dai tempi antichi. Naturalmente, nella vita quotidiana destra e sinistra sono molto differenti l'una dall'altra, ma nella meccanica quantistica la situazione appare diversa: rispetto all'inversione degli assi le funzioni ψ risultano funzioni pari o dispari, hanno cioè una determinata parità intrinseca. Dal momento che la grandezza che di solito si misura è $|\psi|^2$ si è portati a pensare, per estrapolazione, che in tutti i fenomeni fisici le leggi non cambino quando si inverte lo spazio.

È in questo senso che si dice (Wigner) che i fenomeni fisici *conservano la parità*: di conseguenza se un fenomeno fisico avviene nello spazio reale, il fenomeno «visto allo specchio» è un fenomeno perfettamente possibile. E tuttavia: «Negli anni 1954-56 venne fuori un nuovo enigma, chiamato enigma ϑ-τ. Attualmente si sa che i mesoni ϑ e τ sono la stessa particella, che si chiama comunemente K. In quegli anni tuttavia si sapeva soltanto che c'erano delle particelle che si disintegrano in due mesoni π e delle particelle che si disintegrano in tre

mesoni π. Esse venivano chiamate rispettivamente ϑ e τ: τ era il nome dato ad essa da Powell nel 1949. Col tempo le misurazioni divennero piú accurate e la maggiore precisione mise in evidenza sempre piú chiaramente l'esistenza di un enigma. Da un lato era chiaro che ϑ e τ avevano, con grandissima precisione, la stessa massa. Si trovò che si comportavano in modo identico anche sotto altri aspetti. Sembrava cosí che ϑ e τ fossero veramente la stessa particella, che si disintegrava in modi differenti. D'altro lato, esperimenti sempre piú accurati dimostrarono anche che ϑ e τ non avevano la stessa parità e che perciò non potevano essere la stessa particella» [*ibid.*, p. 72].

La soluzione dell'enigma ha richiesto la revisione di una radicata «classificazione naturale» [per questa terminologia, cfr. Bloor 1976]. Scrive ancora Chen Ning Yang [1961]: «Nell'estate del 1956, Tsung Dao Lee e io esaminammo i fondamenti sperimentali allora disponibili di questo concetto, e giungemmo alla conclusione che, contrariamente alla convinzione generale, non esisteva in realtà alcuna prova sperimentale della simmetria tra destra e sinistra per le interazioni deboli. Se la simmetria tra destra e sinistra non è valida per le interazioni deboli, il concetto di parità non è applicabile al meccanismo di decadimento delle particelle ϑ e τ, e perciò esse potevano essere la stessa particella, come ora sappiamo che in realtà è» (trad. it. p. 72).

Si trattava allora di controllare sperimentalmente se la simmetria tra destra e sinistra potesse venir violata per le interazioni deboli. Il primo di questi esperimenti (1956: Wu, Ambler, Hayward, Hoppes, Hudson) consistette sostanzialmente nell'osservazione della distribuzione angolare degli elettroni nel decadimento β di un isotopo del cobalto, il ^{60}Co. I nuclei di cobalto hanno degli spin che, per temperature abbastanza basse, è possibile allineare mediante un forte campo magnetico: si constatò allora che gli elettroni vengono emessi di preferenza nella direzione dell'asse dello spin nucleare, in un verso che per convenzione viene descritto come quello in cui il nucleo ruota in senso sinistrorso (antiorario) attorno alla direzione preferenziale (fig. 6*a*). Ma secondo la stessa convenzione un nucleo apparirebbe ruotare in senso destrorso (orario) se visto in uno specchio (fig. 6*b*). Ciò può venir interpretato in termini di funzioni d'onda, sostenendo appunto che la parità non è conservata. La notizia degli esperimenti venne data in un numero del «New York Times» nel gennaio 1957 e questo fu il commento di Pauli: «Sí, è stata un'esperienza drammatica... Ciò che mi colpisce non è il fatto che 'Dio sia solamente mancino', ma il fatto che nonostante ciò egli si dimostri simmetrico quando si esprime mediante interazioni forti. In breve, il problema è quello di stabilire perché le interazioni forti abbiano una simmetria destra-sinistra. Come può la forza di un'interazione produrre, o creare, gruppi di simmetria, leggi di invarianza e di conservazione? Il problema mi ha portato a soluzioni premature e sbagliate... Non conosco alcuna buona risposta» [citato in Matthews 1971, trad. it. p. 104]. Tali parole bene esemplificano quello stato psicologico diffuso che accompagna un «riorientamento» nella ricerca, quando classificazioni «naturali» (cioè bene radicate in una data tradizione di ricerca) cominciano a «sfocarsi» [Kuhn 1962, trad. it. pp. 111-12].

Ma la «crisi» è benefica. Hermann Weyl [1952] piú volte ripete che in arte

raramente l'asimmetria è solo mancanza di simmetria. L'osservazione calza anche alla fisica.

a) Quando l'anomalia diventa una regola, la ristrutturazione concettuale cambia lo stato di entità relegate ai margini. È il caso del neutrino (cfr. § 3.1). Se si ha una particella dotata di spin, si può considerare uno stato in cui il senso dello spin è sinistrorso rispetto alla direzione del moto. Per una particella dotata di massa questa non è però una situazione ben definita: essa deve viaggiare, infatti, con una velocità minore di quella della luce e la situazione ora descritta si riferisce a un osservatore particolare. È sempre possibile trovare un altro osservatore che supera in velocità la particella, ribaltando in questo modo il verso del suo moto relativo. Ambedue gli osservatori, infine, registrano lo stesso verso di rotazione per lo spin, per cui il secondo osservatore vede la particella che ruota in senso destrorso attorno al verso del moto relativo. Ma per un neutrino, *che ha massa nulla*, le cose vanno diversamente. Per avere energia e quantità di moto diverse da zero esso deve viaggiare con la velocità della luce (che è l'equivalente relativistico di una velocità infinita). Nessun osservatore può allora superare il neutrino, di modo che un neutrino che ruoti in senso antiorario attorno al verso del moto apparirà animato da tale moto per tutti gli osservatori. È perciò perfettamente coerente con la relatività supporre che in natura esistano solamente neutrini sinistrorsi e che le interazioni deboli siano vincolate da questa restrizione. Evidentemente ciò non è conciliabile con l'invarianza rispetto alle riflessioni poiché nel «mondo riflesso» i neutrini sarebbero solamente destrorsi.

b) «Il punto importante è che, se si cambia la definizione di riflessione speculare, la simmetria per riflessione speculare può essere ristabilita» [Chen Ning Yang 1961, trad. it. p. 76]. Alla usuale riflessione speculare si affianca il cambiamento delle particelle nelle rispettive antiparticelle (cfr. § 2.1). Per esempio, se

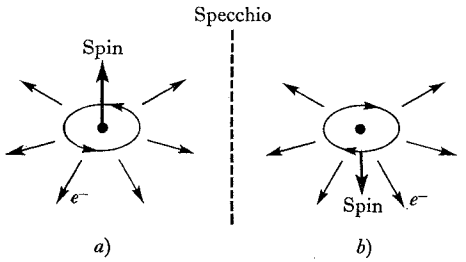


Figura 6.

Esperienza di Wu e collaboratori. Introducendo il preparato radioattivo in un ambiente a bassa temperatura (verso lo zero assoluto) e sottoponendolo a un campo magnetico molto intenso si ottiene questa condizione: praticamente tutti i momenti magnetici, e quindi gli spin, dei nuclei di Co sono orientati nella direzione del campo, poiché la bassissima temperatura elimina le fluttuazioni termiche. Gli elettroni provenienti dal decadimento sono preferenzialmente emessi dalla parte indicata nel disegno in *a*), cioè in direzione opposta a quella degli spin dei nuclei. Tuttavia nell'immagine speculare *b*) del fenomeno cambia il senso dello spin, ma non cambia la direzione di emissione degli elettroni. (Da Toraldo di Francia 1976).

nella figura 6*b* si cambia il nucleo nel suo antinucleo, quest'ultimo risulta possedere carica negativa e pertanto, se ruota in senso inverso al nucleo della figura 6*a*, avrà momento magnetico orientato nella medesima direzione. Questa mossa è nota (Wigner) come coniugazione della carica: «Il principio di ragion sufficiente non è in contraddizione con l'esperienza sul cobalto (e con tutte le altre esperienze sulle interazioni deboli), in quanto esiste l'antimateria. Grazie alle interazioni deboli possiamo ora dare un significato assoluto ai termini 'destra' e 'sinistra'. Tuttavia i due significati vengono scambiati se passiamo dal nostro mondo, formato di materia, a un antimondo, anch'esso concepibile, formato di antimateria. In questo senso, *l'esistenza dell'antimateria rappresenta la modalità scelta dalla natura per mantenere viva una forma ridotta della simmetria sinistra-destra.* Dopo una prima sorpresa, questa soluzione può ancora soddisfare il nostro desiderio di simmetria. Ma la storia non è ancora finita. Nel 1964 si è scoperto che una nuova interazione superdebole influenzava certi modi di disintegrazione dei kaoni [J. H. Christenson, J. W. Cronin, V. R. Fitch, R. Turlay] e questi nuovi fenomeni non sono simmetrici, nemmeno quando sono osservati attraverso lo specchio che trasforma le particelle in antiparticelle. Tuttavia la simmetria può venir ristabilita utilizzando uno specchio che non solo inverte la sinistra e la destra e trasforma particelle in antiparticelle, ma che inverte pure la "freccia del tempo"» [Amaldi 1979, p. 94].

Si procede, dunque, per slittamenti successivi, in modo da lasciare sempre uno spiraglio a quello che si potrebbe scherzosamente definire «il fascino discreto» delle simmetrie. Si tratta ora di dilatare la nozione originale di riflessione speculare in modo da includervi anche *l'inversione temporale.* Ora, dai principî della meccanica quantistica relativistica si può ottenere un teorema che afferma sostanzialmente che, se a un dato processo fisico si applicano, in un ordine qualsiasi, le tre trasformazioni materia-antimateria (cioè coniugazione della carica C), inversione nello spazio (parità P) e inversione del tempo (T) si ottiene un processo ammissibile. In altri termini: tutti i processi fisici devono essere invarianti per trasformazione CPT (Jost, 1958). Ora, «con grandissima probabilità... si può ritenere valida l'inversione temporale in tutti i processi. Se è valida l'invarianza rispetto a T, e si deve ancora avere per il teorema citato l'invarianza rispetto a CPT, ne deriva che si deve sempre avere l'invarianza CP. Ma CP appare violata e nella stessa misura bisogna ammettere che lo sia T. E allora?... Queste questioni sono ancora molto misteriose» [Toraldo di Francia 1976, pp. 473-74].

Ma, a conclusione di questo e del precedente paragrafo, si può almeno affermare che la sistematica classificazione delle particelle riceve da quest'ordine d'idee un ulteriore contributo: le interazioni forti sono vincolate da ben dodici leggi di conservazione: energia; quantità di moto; momento angolare; carica; numero elettronico; numero muonico; numero barionico; inversione temporale (T); inversione spaziale e coniugazione di carica combinate (CP); inversione spaziale da sola (P) e coniugazione di carica da sola (C); stranezza; spin isotopico. L'ordine di questo elenco non è casuale: scendendo la scala delle intensità (cfr. § 3.3) la conservazione dello spin isotopico viene perduta già nel primo gradino, le interazioni elettromagnetiche. Con le interazioni deboli anche la conservazione della stranezza, la conservazione della parità e la conservazione della coniugazione di carica vanno perdute, ma resta valida la conservazione «combinata» CP. La «questione molto misteriosa» potrebbe costituire, in futuro, un primo rilevante indizio del fatto che esiste nella scala delle interazioni un gradino «intermedio» in cui l'invarianza CP e l'invarianza T sono violate.

3.4. «Legge e ordine» tra gli adroni.

«Nel 1910 il matematico Oswald Veblen ed il fisico James Jeans stavano discutendo la riforma dell'insegnamento della matematica all'Università di Princeton. "Possiamo benissimo abolire la teoria dei gruppi", arguiva Jeans, "è una materia che in fisica non servirà mai a niente"... Fu un fatto di una certa importanza per la storia della scienza a Princeton che Veblen non abbia seguito il consiglio di Jeans. Per ironia del fato, recentemente la teoria dei gruppi è diventata uno dei temi centrali della fisica, ed ora essa domina il pensiero di tutti coloro che cercano di capire le particelle fondamentali della natura e, vedi caso, i due pionieri dell'impiego della teoria dei gruppi alla fisica, Hermann Weyl e Eugene P. Wigner, erano entrambi professori a Princeton» [Dyson 1964, trad. it. p. 114].

Gli aspetti piú propriamente matematici (gruppi e loro rappresentazioni) sono trattati in altre voci della presente *Enciclopedia* (in particolare «Applicazioni» e «Invariante»). Qui sarà opportuno ricordare che la potenza che la teoria dei gruppi rivela nella fisica sembra dipendere, sostanzialmente, dal fatto che, ogni volta che un «oggetto» fisico presenta delle simmetrie, in base alle leggi della meccanica quantistica deve esistere un gruppo G di operazioni compatibili con tali simmetrie, le rappresentazioni del quale corrispondono esattamente ai possibili stati quantici dell'oggetto considerato.

Esemplificando: si consideri un atomo che si muove in un gas rarefatto. Per esso tutte le direzioni dello spazio sono equivalenti e quindi esso presenta la simmetria dell'ordinario gruppo R_3 delle rotazioni. Tra le rappresentazioni di R_3 c'è quella ternaria, che corrisponde agli stati dell'atomo che hanno spin uguale a 1. Essi si presentano sempre in gruppi di tre, con la medesima energia. S'immagini ora un campo magnetico in grado di distruggere la simmetria sferica del fenomeno: le tre energie uguali vengono ora a separarsi un po' e i tre stati corrispondenti possono essere rivelati dallo spettroscopio come una terna visibile di linee spettrali.

Si esamini ora una situazione apparentemente differente. Come si è detto ai §§ 3.1 e 3.2, i tre pioni hanno approssimativamente la stessa massa e le stesse interazioni nucleari. Se s'immagina che essi costituiscano una rappresentazione ternaria di un gruppo G, molte loro proprietà potranno venir ricavate senza aver la piú lontana idea circa la natura degli elementi di G. Nella fattispecie il gruppo G è qui il gruppo $SU(2)$ o gruppo dello spin isotopico (cfr. § 3.2) ed è rilevante notare che certe «anticipazioni» delle proprietà dei pioni vennero ricavate da Nicholas Kemmer in base a considerazioni astratte nel 1938, nove anni prima della scoperta del primo pione. La cosa interessante è, però, che «$SU(2)$ e R_3 costituiscono essenzialmente lo stesso gruppo, se ci si interessa solo delle pro-

prietà algebriche delle operazioni gruppali, senza preoccuparsi della differenza delle rappresentazioni matriciali» [Omnès 1970, p. 62].

Ma si venga ora alle interazioni forti. Il gruppo G cercato «dovrà includere il gruppo dello spin isotopico come sottogruppo, dal momento che esso è conservato dall'insieme di tutte le interazioni forti. La maniera in cui questo gruppo ingloberà la stranezza può realizzarsi in due modi: o la stranezza è conservata separatamente senza alcuna relazione con il resto delle leggi di conservazione – come del resto è sembrato fino quasi ai giorni nostri – o il gruppo d'invarianza delle interazioni piú forti ingloba la conservazione della stranezza come lo spin isotopico ingloba la conservazione della carica – e qui la terza componente dello spin isotopico T_3 prende essenzialmente il posto della carica elettrica. Quest'ultimo caso è il piú interessante a priori, in quanto conduce a una assimilazione della stranezza alla carica elettrica e, quindi, a una maggior unità e alla speranza di una maggiore semplicità. È questo caso dunque che si esaminerà qui» [*ibid.*, p. 340]. Una questione preliminare sta nel decidere se si cerca un gruppo G le cui rappresentazioni unitarie siano di dimensione finita oppure infinita, cioè se G sia compatto o no (fisicamente: se le particelle debbano venir classificate mediante multipletti in cui compare un numero finito oppure infinito di particelle). R_3 è un gruppo compatto: altri gruppi impiegati ampiamente nella fisica (per esempio il gruppo di Poincaré-Lorentz) non lo sono. Ora «anche se la classificazione di una infinità di particelle è degna di considerazione, ci si occuperà nel seguito del caso dei multipletti finiti, cioè del caso di un gruppo G compatto. Il gruppo G deve permettere l'esistenza almeno di due quantità additive simultaneamente misurabili che corrispondono rispettivamente alla stranezza e alla carica... Anzi, restringeremo la generalità e affronteremo il caso piú semplice, quello in cui G ingloba solo due quantità additive simultaneamente misurabili... Il gruppo G, che ha un certo numero di generatori, avrà dunque due generatori corrispondenti a T_3 e Y (Y è legato alla stranezza) che commutano

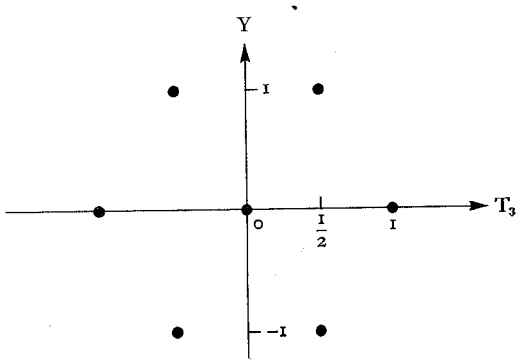


Figura 7.

Per descrivere le rappresentazioni di $SU(3)$ si adotta un sistema di coordinate M e T_3, ove T_3 è la terza componente dello spin isotopico e, per semplificare, si pone $Y = 2M/\sqrt{3}$. La figura descrive, in particolare, la rappresentazione {8}. (Da Omnès 1970).

tra loro, mentre nessun altro generatore commuta simultaneamente con T_3 e Y. In termini matematici si dice che un siffatto gruppo è di rango 2. All'interno di una rappresentazione unitaria irriducibile, gli stati sono distinti in base al valore di due numeri quantici additivi. Per un confronto: il gruppo delle rotazioni R_3 è di rango 1: all'interno di una rappresentazione irriducibile, gli stati sono determinati dall'assegnazione di un solo numero quantico additivo, proiezione del momento angolare sull'asse z. È una circostanza molto felice che tutti i gruppi compatti siano stati classificati da Elie Cartan e che non esista che un numero ristretto che soddisfa alle condizioni che abbiamo imposto. Possiamo quindi passare tali gruppi in rivista, al fine di determinare se uno qualsiasi di essi svolga un ruolo nella fisica. In particolare il gruppo "per bene", se esiste, dovrà possedere delle rappresentazioni unitarie irriducibili, nelle quali potremo classificare particelle che presenteranno tutte gli stessi numeri quantici cinematici (spin e parità) e lo stesso numero barionico» [*ibid.*, p. 341].

Fatta la «rassegna», un solo gruppo risulta il candidato adeguato per G: il gruppo $SU(3)$, proposto per la prima volta da Gell-Mann e Ne'eman. $SU(3)$ è il gruppo delle matrici 3×3 unitarie e unimodulari, cioè delle matrici 3×3 tali che $UU^+ = I$ e $\det U = 1$. Il nome del gruppo ricorda semplicemente queste proprietà: la U in $SU(3)$ significa unitario; la S significa speciale, cioè la condizione di unimodularità delle matrici del gruppo; la cifra 3, infine, specifica evidentemente che si tratta di matrici di 3 righe e 3 colonne. $SU(3)$ è una struttura matematica piú complicata di $SU(2)$: è quindi piú difficile associarvi una modellizzazione geometricamente intuitiva come si è fatto, poco piú sopra, movendo dalle rotazioni R_3 dello spazio ordinario. (Ciò però è possibile [cfr. Dyson 1964, trad. it. pp. 128-32]).

La rappresentazione {8}, cioè a otto dimensioni, di $SU(3)$ (cfr. fig. 7), è stata battezzata da Gell-Mann «ottuplice sentiero». Questa rappresentazione

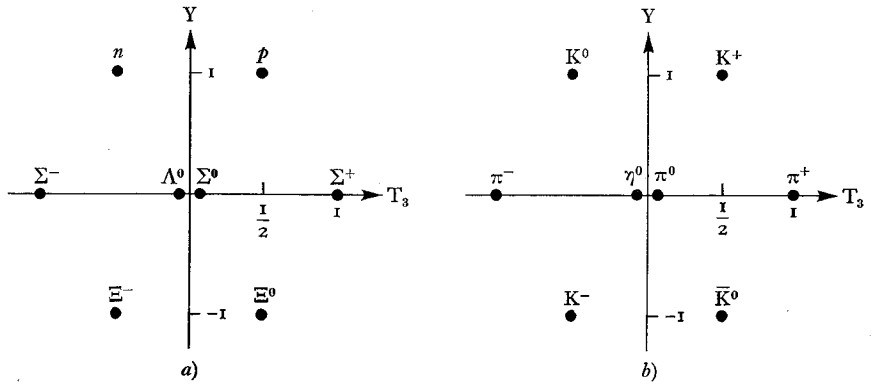


Figura 8.

a) La rappresentazione {8} dei barioni. *b*) La rappresentazione {8} dei mesoni, in cui compare anche il cosiddetto mesone η. (Da Omnès 1970).

contiene uno scalare isotopico $T = 0$, $Y = 0$; due spinori isotopici $T = 1/2$, $Y = +1$; un vettore isotopico $T = 1$, $Y = 0$. Essi individuano un singoletto, due doppietti e un tripletto (per questa terminologia cfr. § 3.2) e possono quindi venir interpretati in termini di barioni (cfr. fig. 8*a*) e di mesoni (fig. 8*b*).

In altre rappresentazioni di $SU(3)$ vi sono dieci, ventisette elementi o anche piú. Per primo Gell-Mann ha notato che la simmetria di un insieme di dieci elementi corrispondeva esattamente a un insieme di nove tra i barioni conosciuti, completato da un singoletto ancora mancante, che egli ha battezzato, in anticipo, Ω^-. I membri conosciuti di questo insieme di dieci elementi erano una quaterna Δ, un'altra terna Σ e un'altra coppia Ξ. Il singoletto previsto è stato scoperto il 31 gennaio 1964 al Brookhaven National Laboratory (Long Island, New York), su fotografie prese in una camera a bolle [cfr. Flower e Samios 1964]. E inoltre: «Vi sono oggi indizi schiaccianti che il comportamento delle particelle a interazioni forti effettivamente corrisponda alla simmetria astratta descritta dal gruppo $SU(3)$; questa simmetria non è perfetta, perché è rotta da una perturbazione relativamente debole, che riduce il gruppo $SU(3)$ a [un] suo sottogruppo $U(2)$; la simmetria che resta, descritta dal gruppo $U(2)$, è essenzialmente identica a quella del gruppo dello spin isotopico... L'intero quadro delle particelle ad interazioni forti, mercé l'intervento di queste semplici ed efficaci concezioni di teoria dei gruppi, è stato trasformato: dal caos [cfr. quanto detto al § 3.2] a un ordine notevole» [Dyson 1964, trad. it. p. 132].

3.5. Democrazia o aristocrazia? Ipotesi del *bootstrap*.

Nell'articolo «Fisica» di questa stessa *Enciclopedia* (in particolare al § 4.1) due atteggiamenti di pensiero differenti – ma entrambi notevolmente fecondi – sono simbolizzati dai nomi di Democrito e di Pitagora. Democrito significa, nel contesto della fisica delle particelle, l'appello ai costituenti ultimi; Pitagora al potere risolutivo della matematica. Qui non si traccerà minimamente un bilancio delle varie strutture matematiche impiegate oggi per le particelle elementari – per cui il lettore è rimandato ad altri articoli della presente *Enciclopedia*. Può apparire naturale da quanto detto nel § 3.4 una riconsiderazione del passaggio dal «caos» dei §§ 3.1 e 3.2 all'«ordine» del § 3.4 come una vittoria dell'atteggiamento intellettuale che vede in strutture matematiche particolarmente eleganti e semplici «l'essenza primordiale dell'universo fisico» (cfr. anche quanto detto nell'articolo «Numero» di questa stessa *Enciclopedia*, vol. IX, p. 944). Ma l'approccio *via* teoria dei gruppi, almeno come è stato fin qui delineato, resta al di qua di una vera e propria teoria *dinamica* degli adroni.

La prima grande risposta a quest'ultima esigenza è stato il programma promosso da Geoffrey Chew verso il 1960. «È necessario sottolineare la novità di questo suo approccio in rapporto alla tradizione filosofica e scientifica dell'Occidente. Dall'epoca degli antichi Greci, abbiamo tentato di comprendere non solo la materia ma anche la totalità del mondo reale riducendola a delle componenti elementari. La scienza moderna si è interessata agli strati diversi della materia: atomi prima, poi nuclei ed elettroni, infine protoni e neutroni. Verso la fine degli anni '50, l'esperienza del passato e la constatata proliferazione di particelle dette "elementari" sembravano suggerire l'esistenza di un nuovo strato di sottostrutture. Il tentativo di fermare questo regresso apparentemente infinito, che si è avuto in quell'epoca, ha dunque rappresentato un fatto importante nella storia della disciplina» [Amaldi 1979, pp. 84-86].

L'evidenza sperimentale indica che, a patto di mettere in gioco l'energia cinetica E_c sufficiente, è possibile in un urto avere trasformazioni reversibili di certe particelle (per esempio A) in altre particelle (per esempio B).

La particella A non ha quindi «maggiori diritti di elementarità» [Toraldo di Francia 1976, p. 476] della particella B. «In questa visione il mondo risulta, pertanto, composto *non* di particelle, ma di un *complesso di entità* che noi vediamo *granularizzarsi* in ciò che chiamiamo *particelle*. Ciascuna particella, poi, per quanto visto nella teoria dei campi, ha la possibilità di emettere *virtualmente* tutte le particelle con cui essa può interagire. Un protone, per esempio, emetterà e assorbirà virtualmente tutta la serie di particelle fortemente interagenti (pioni, kaoni, particelle Λ_0, ecc.)» [*ibid.*].

Chew [1965] ha battezzato «democrazia nucleare» un regime in cui tutte le particelle che subiscono l'interazione forte si situano «su una stessa base di equivalenza dinamica»: è l'ipotesi detta *dynamical bootstrap* (Chew, Maldestam) per significare entità che sostengono se stesse (*bootstrap* significa 'laccio da scarpe').

«Questo approccio era nuovo, dal momento che non impiegava dei campi fondamentali..., ma era evidentemente limitato in quanto poneva problemi matematici insolubili e non spiegava l'esistenza dei leptoni e dei fotoni che non subiscono interazioni forti e che certo sono incapaci di generarsi per *bootstrap*. Tuttavia aveva il grande merito di porre in una luce del tutto nuova le questioni filosofiche concernenti i blocchi costitutivi della materia. Ciò premesso, dopo la scoperta di un nuovo strato di sottostrutture, i fisici avrebbero il dovere di considerare almeno tre alternative invece che due: I) le sottostrutture sono 'elementari'; II) le sottostrutture sono formazioni composite di altri blocchi; III) le sottostrutture sono degli stati legati autocoerenti di se stesse» [Amaldi 1979, p. 86]. Gli stessi proponenti non si sono nascoste le difficoltà del programma, e, nemmeno, il rischio di aver instaurato però una democrazia solo apparente, incapace cioè di escludere l'eventualità di particelle dotate di uno statuto «elementare» privilegiato. E «per quanto ne sappiamo oggi, l'ipotesi del *bootstrap non* si applica al mondo degli adroni e degli aristocratici continuano a dominare in questo strato della materia. Sono detti "quark"» [*ibid.*, p. 87].

3.6. Quark e antiquark.

«Al momento attuale sembra che si possa considerare l'invarianza relativamente a $SU(3)$ come una ben stabilita proprietà delle interazioni forti... Una delle ipotesi piú ardite e piú interessanti consiste allora nel supporre che esistano della particelle, i *quark*, che si trasformano rispetto a $SU(3)$ come gli elementi della rappresentazione $\{3\}$» [Omnès 1970, p. 359], cioè di una particolare rappresentazione di $SU(3)$ (cfr. fig. 9*a*). Formalmente non è poi difficile mostrare

che tutte le rappresentazioni di $SU(3)$ possono venir costruite mediante prodotti diretti delle rappresentazioni $\{3\}$ e $\{3^*\}$ (rappresentazione *coniugata* della precedente: cfr. fig. 9*b*): si potranno allora concepire tutte le particelle come stati legati di quark e di antiquark.

L'ipotesi dei quark (e antiquark), proposta indipendentemente da Gell-Mann e Zweig nel 1963 (il nome è tratto da una frase di *Finnegan's Wake* di Joyce, «Three quarks for mister Mark») ha senza dubbio l'ambizione di mettere ordine nello zoo degli adroni spiegando la «tavola periodica subnucleare» del § 3.1 proprio come la congettura di Bohr ha spiegato i raggruppamenti degli elementi nella tavola periodica di Mendeleev. Cosí, i barioni sono formati ciascuno da tre quark; i mesoni sono formati ciascuno da un quark e da un antiquark (cfr. del resto quanto è detto su quark e antiquark nell'articolo «Materia» in questa stessa *Enciclopedia*, vol. VIII, pp. 918 sgg.).

Ma i quark, tanto per riprendere la problematica del § 1.3, esistono o sono finzioni? Per verificarne l'esistenza (nel senso in cui s'intende questa locuzione nel presente articolo, e cioè con le qualificazioni del § 1.3) deve essere possibile «osservarli» come entità separate, per esempio mediante le tracce che essi lasciano in una camera a bolle; queste dovrebbero rivelare la massa e la quantità di moto dei quark attraverso la curvatura delle tracce in un campo magnetico e le restrizioni conseguenti alle leggi di conservazione dell'energia e della quantità di moto. Anche la loro carica dovrebbe essere rivelata in base alla densità delle bollicine lungo la traccia; per particelle che viaggiano con velocità vicina a quella della luce, questa densità dipende dal quadrato della carica, in modo che la traccia di una particella avente carica 1/3 dovrebbe essere di un ordine di grandezza

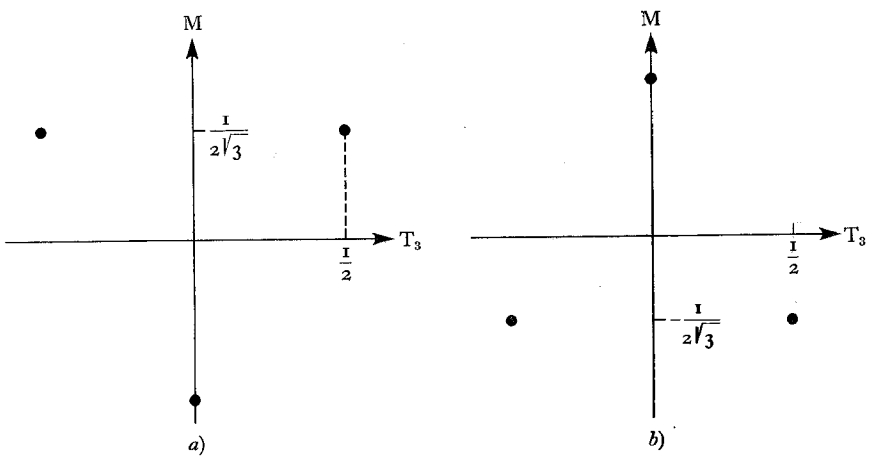


Figura 9.

a) La rappresentazione $\{3\}$ di $SU(3)$ descritta con un diagramma analogo a quello della figura 7. ***b***) La rappresentazione coniugata $\{3^*\}$ è ottenuta dalla rappresentazione $\{3\}$ considerando le matrici complesse coniugate α^* delle matrici della rappresentazione $\{3\}$.

piú debole di quella di particelle aventi una carica normale e dovrebbe quindi essere facilmente rivelabile. Se gli adroni sono composti di quark, questi ultimi dovrebbero essere prodotti negli urti adronici, ammesso che il processo sia compatibile con tutte le leggi di conservazione pertinenti. A causa della loro carica elettrica frazionaria (si veda la tabella 3 nel citato articolo «Materia») non c'è però alcuna possibilità di produrne uno alla volta, poiché uno stato finale che implichi particelle con carica elettrica 1/3 di quella dell'elettrone non può verificarsi in urti conservanti la carica fra particelle normali con cariche intere; i quark potrebbero però essere prodotti in terzetti o in coppie quark-antiquark. Queste ultime avrebbero tutte le cariche totali nulle – elettrica, barionica, ipercarica – di modo che, per ciò che concerne le leggi di conservazione, una coppia di quark potrebbe essere prodotta in base a condizioni iniziali molto generali. Poiché la differenza fra le cariche dei quark – come si evince subito dalla tabella – è intera, il quark piú pesante, quello s, di carica $-1/3$ può decadere in base alla normale interazione debole nel quark u di carica 2/3 con emissione di leptoni: $s^{-1/3} \rightarrow u^{2/3} + e^{-1} + \bar{\nu}_e^0$. Questo è lo stesso meccanismo di decadimento β in base al quale un neutrone decade in un protone; il quark piú leggero deve essere però stabile in un ambiente composto di particelle a carica intera normale, perché non c'è alcun modo coerente con la conservazione della carica in base a cui esso può sbarazzarsi della sua carica elettrica frazionaria. Perciò una volta che il quark sia stato «prodotto», c'è da attendersi che la sua presenza possa essere «rivelata» in qualche modo e «resa visibile».

Dunque i quark «avrebbero proprietà straordinarie che li farebbero riconoscere sicuramente; per esempio avrebbero cariche elettriche 1/3 o 2/3 della carica dell'elettrone! Naturalmente i fisici li hanno cercati con cura un po' da per tutto, ma finora senza successo» [Segrè 1976, p. 256]. D'altro canto, «l'aspetto piú rivoluzionario del modello dei quark concerne appunto la proprietà secondo cui i quark possiedono cariche elettriche frazionarie. Tutte le particelle note, infatti, hanno una carica che è multipla della carica elettrica dell'elettrone e_0 e i fisici sono riusciti a immaginare formalismi in cui questo può apparire come un fatto naturale. Un numero barionico e una ipercarica frazionaria sembrano ugualmente poco naturali, ma meno rivoluzionari. In ogni modo, fin dal suo debutto, il modello dei quark ha conosciuto successi impressionanti nella predizione delle proprietà statiche degli adroni. Eppure la stragrande maggioranza dei ricercatori hanno assai esitato prima di accettare i quark come costituenti elementari della materia adronica (prima di accettare, cioè, che i campi dei quark siano le entità fondamentali)» [Amaldi 1979, pp. 87-88].

4. *Approfondimenti.*

La determinazione delle proprietà delle particelle cui si è accennato nei paragrafi precedenti si basa sullo studio delle loro interazioni con campi esterni, oppure – o anche – con altre particelle. Un riferimento indirizzato a questi scopi richiede in genere: 1) una sorgente di particelle (dato che i processi di accelera-

zione si basano sull'azione di campi elettrici e magnetici, la sorgente dovrà fornire particelle elettricamente cariche, cioè ioni); 2) un apparato acceleratore; 3) campi elettrici e magnetici esterni; 4) un bersaglio materiale col quale la particella studiata possa interagire; 5) uno o piú rivelatori.

L'energia raggiungibile dall'acceleratore è essenziale nella definizione delle esperienze; quanto piú elevata è questa, tanto piú vasto è il campo esplorabile. Infatti, se si studiano processi di produzione, per creare una particella di massa m occorre un urto nel quale sia disponibile un'energia $E \geq mc^2$. Tanto piú alta è la massa m, tanto maggiore l'energia necessaria.

Sotto questo profilo, la fisica delle particelle è davvero una «fisica delle alte energie». E, benché le sorgenti naturali – come gli elementi radioattivi o i raggi cosmici – abbiano svolto un ruolo rilevantissimo nello sviluppo di questa disciplina, oggi non sono competitivi con le sorgenti artificiali. In particolare va ribadito che «se oggi possiamo parlare di unificazione delle interazioni e delle sottostrutture nei nucleoni, è solo perché, nel corso dei cinquanta ultimi anni gli acceleratori di particelle, che all'origine erano dei piccoli strumenti di lavoro artigianale, sono divenuti degli strumenti grandi e complessi. Lo stato delle nostre conoscenze è modellato non solo dalla meccanica quantistica e dalla teoria dei gruppi ma anche dalle possibilità tecniche e dalle limitazioni degli acceleratori e dei rivelatori» [*ibid.*, p. 130]. Non è possibile, in questa sede, dare un quadro degli apparati sperimentali e della loro evoluzione né tracciare in breve una storia dei vari acceleratori (elettrostatici, ciclici, lineari, ecc.) e rivelatori (contatori a scintillazione, contatori a cristallo, camere a scintilla, camere a fili, camere a bolle, emulsioni nucleari, ecc.). Il piú volte citato saggio di Amaldi fornisce esemplarmente piú di una prova del carattere «complementare» della evoluzione delle idee e del progresso di queste strutture sperimentali.

Va aggiunto ancora che uno dei grandi temi unificanti delle ricerche sulle particelle elementari – su cui si intrecciano aspetti teorici e realizzazioni sperimentali – è, senza alcun dubbio, quello dell'unificazione delle interazioni fondamentali. Su questo tema il lettore troverà non poche indicazioni in altre voci della presente *Enciclopedia* (quali «Forza/campo», «Interazione», ecc.); in questa sede si è insistito di piú sulle differenti vesti che ha assunto il problema dei costituenti della materia e sulle revisioni concettuali che questi slittamenti del problema hanno via via favorito o imposto. La scelta di alcuni ulteriori approfondimenti – che non pretende ad alcuna sistematicità (tanto piú che alcune delle teorie cui si farà riferimento nelle pagine successive sono vivacemente impugnate) – si propone di dare semplicemente un'idea del carattere estremamente «mobile» che attualmente ha la ricerca, cosí compenetrata di strutture teoriche e di poderosi apparati sperimentali.

4.1. Partoni e quark.

Si è già detto (cfr. § 3.6) che non sono mai stati «osservati» quark isolati (sono stati però «osservati» stati legati di quark, cioè particelle instabili decadenti in modi e con vite medie previste dalla teoria). Esiste un'altra serie di esperimenti che però può dare informazioni sull'argomento: le reazioni ad alta energia fra elettrone e nucleone e fra neutrini e nucleoni. Il principio su cui si basa la discussione di questi esperimenti è analogo a quello sul quale si fonda l'analisi dell'urto fra atomi e particelle nel celebre esperimento di Rutherford. L'atomo di Thomson era, in breve, un sistema nel quale la carica elettrica positiva era distribuita con continuità su tutto il volume dell'atomo. I risultati di Rutherford mostrarono invece che la carica positiva era concentrata in un nucleo centrale assai piccolo che racchiudeva praticamente anche tutta la massa. In maniera analoga gli esperimenti d'urto con elettroni e neutrini di alta energia sembrano indicare che nell'interno dei nucleoni l'urto avviene non contro una distribuzione estesa, ma contro una distribuzione di corpuscoli puntiformi che interagiscono sia con gli elettroni sia con i neutrini. Richard Feynman ha introdotto il termine 'partoni' per tali sottostrutture proprio per evitare frettolose identificazioni con gli ipotetici quark. Ma quando si è cominciato – agli inizi degli anni '70 – a confrontare i dati delle esperienze con elettroni (Stanford) con quelli delle esperienze con i neutrini (Cern, Ginevra) l'idea che partoni e quark siano la stessa cosa ha preso maggior vigore e autorità. Infatti: dalle esperienze con elettroni si deve concludere che i partoni hanno spin 1/2. Dal confronto fra la diffusione degli elettroni e quella dei neutrini da parte di protoni e neutroni si ricavano le cariche elettriche dei partoni: gli elettroni *vedono* le cariche elettriche, i neutrini non le vedono; questi ultimi interagiscono soltanto attraverso accoppiamenti deboli. Il confronto sperimentale indica per dette cariche elettriche i valori di 2/3 e 1/3. Non si dà il segno in quanto l'accoppiamento dipende dal quadrato delle cariche.

C'è quindi una evidenza sperimentale largamente in favore dell'ipotesi dei quark come costituenti fondamentali della materia. Ma la struttura del nucleone è piú complessa ancora: ammette, in aggiunta ai partoni-quark, qualcosa d'altro, i quanti di quelle forze quarkiche che tengono i quark confinati nel volume di un nucleone. Si tornerà tra poche pagine su questo argomento.

4.2. Cromodinamica quantistica.

In questi ultimi anni lo studio dei quark mediante esperienze di diffusione è stato approfondito grazie al rapido sviluppo di un nuovo apparato della fisica delle alte energie, gli anelli di accumulazione: si tratta, sostanzialmente, di magneti anulari in cui possono circolare particelle di data quantità di moto. Se il vuoto è sufficientemente alto, queste particelle potranno circolare liberamente, senza urtare contro molecole presenti, il che li disperderebbe.

L'energia disponibile nella collisione di una particella di energia E con una particella in riposo di massa m_T è data dalla relazione approssimata: (energia nel baricentro) $\simeq \sqrt{2Em_Tc^2}$. Nel baricentro della collisione tra un protone di 20 GeV e un nucleone in riposo ($m_Tc^2 = 0{,}94$ GeV) l'energia è solo 6,2 GeV; nella collisione tra un positrone di 20 GeV e un elettrone in riposo ($m_Tc^2 = 0{,}0005$ GeV) l'energia disponibile si riduce a 0,14 GeV. «Per parecchi anni i fisici hanno sognato di fare un miglior uso dell'energia cosí faticosamente comunicata alle

loro particelle facendo entrare in collisione frontale due fasci uguali di particelle; in questo caso il baricentro della collisione è in riposo nel laboratorio e l'energia disponibile è uguale a due volte l'energia di ciascun fascio. Tuttavia è solo all'inizio degli anni '60 che diversi progressi nella tecnologia degli acceleratori hanno aperto la via alla realizzazione di anelli di accumulazione capaci di conservare due fasci di intensità sufficiente per periodi lunghi abbastanza da produrre un numero non trascurabile di collisioni al secondo in ciascuno dei punti di intersezione dei due fasci. Al Cern la costruzione degli anelli di accumulazione a intersezioni è stata terminata nel 1970 e le collisioni tra due fasci di protoni di 30 GeV hanno rivelato numerosi aspetti inaspettati delle interazioni forti. Ancora piú feconda è stata la via preconizzata da Bruno Touschek che nel 1959 aveva proposto di studiare l'annichilazione di positroni ed elettroni accumulati nello stesso anello magnetico, con l'argomento che preferiva questa reazione alla collisione tra protoni, poiché essa è piú facile da interpretare» [Amaldi 1979, p. 113]. L'anello AdA [per la ricostruzione della sua storia cfr. Bernardini 1978] è stato l'anello di accumulazione elettrone-positrone che ha mostrato la via a tutta una famiglia di macchine riuscite (macchine Vepp a Novosibirsk, Aco ad Orsay, Adone a Frascati, ecc.).

I fenomeni elementari che intervengono nel processo e^+e^- sono illustrati nella figura 10. L'elettrone e il positrone portano ciascuno una energia E e, annichilandosi, producono un fotone di energia $2E$ che è in riposo nel laboratorio, in quanto all'inizio elettrone e positrone avevano uguali velocità in valore assoluto, ma opposte. Il fotone è virtuale: esso può «vivere» un tempo dell'ordine di $\hbar/2E$. Questa breve vita non permette la produzione di adroni normali (come pioni, kaoni o protoni), ma non influisce sulla produzione di leptoni, nella fattispecie di coppie di muoni.

Giova a questo punto introdurre il numero R, definito come il rapporto tra il numero degli eventi in cui si producono adroni e quello in cui si producono delle coppie $\mu^+\mu^-$. Prima delle esperienze era stato previsto, per circa 2 GeV, un valore di R inferiore a 1/10: i risultati ottenuti ad Adone hanno invece fornito per circa 2 GeV un valore di R vicino a 2. Si può interpretare un esito del genere affermando che il fotone virtuale non produce degli adroni, ma coppie quark-

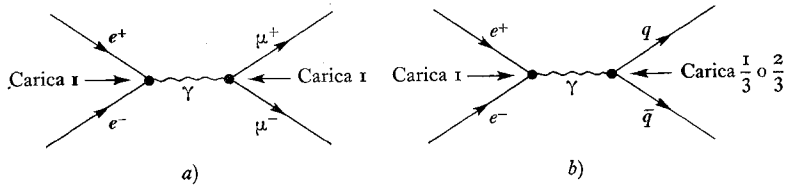


Figura 10.

Nell'annichilazione di un elettrone con un positrone si produce un fotone virtuale di grande energia. La probabilità della produzione di una coppia di muoni (*a*) è proporzionale al quadrato della carica (cioè è proporzionale a 1) mentre la probabilità della produzione di una coppia quark-antiquark (*b*) è proporzionale a $(1/3)^2$ o a $(2/3)^2$ in funzione del valore della carica. (Da Amaldi 1979).

antiquark che si comportano come delle coppie $\mu^+\mu^-$ e si frammentano infine in adroni. Ma questa affermazione non si accorda ancora sotto il profilo quantitativo con i risultati sperimentali: le coppie di quark che possono venir create sono (cfr. tab. 2) $u\bar{u}$, $d\bar{d}$ e $s\bar{s}$ e le loro cariche elettriche sono uguali a 2/3, 1/3 e 1/3 della carica di un muone. Dato che la possibilità di produzione è proporzionale al quadrato della carica, il rapporto R dovrebbe risultare uguale a $(2/3)^2 + (1/3)^2 + (1/3)^2 = 2/3$ e non uguale a 2.

Un tipico *puzzle* per il modello dei quark, che però si può brillantemente risolvere col raffinamento introdotto (1964) da Greenberg per spiegare *altre* difficoltà della teoria (un rendiconto della quale il lettore troverà nell'articolo «Materia» piú volte citato, alle pp. 920-21). In questa prospettiva ogni varietà o «sapore» (*flavour*) dei quark ha un «colore» (*colour*) che può esistere in tre tinte («rosso», «verde» e «blu»). Analogamente, per gli antiquark si definiscono i tre relativi «anticolori».

L'ipotesi dei colori moltiplica per tre il valore $R = 2/3$ di cui sopra, ripristinando in tal modo l'accordo con l'esperienza. Inoltre l'introduzione dei colori apre la via a una spiegazione piú profonda delle forze forti e fornisce una plausibile giustificazione della ragione per cui non si sono «osservati» ancora dei quark liberi.

Le anomalie vengono cosí riassorbite dal modello dei quark. Ma si apre una questione di fondo: perché tutti gli adroni sono incolori? Se la teoria dei «colori» è qualcosa di piú di una struttura formale costruita sulla base del fatto che i barioni rispetto ai colori sono identici (in quanto li contengono tutti e tre) e nei mesoni il colore di un quark è annullato dall'anticolore dell'antiquark, occorre fornire al quesito una risposta che permetta di precisare in che senso il colore è la sorgente del campo che tiene uniti i quark in un adrone. La risposta è di nuovo in una condizione d'invarianza. Il campo elettromagnetico, notoriamente, possiede un'invarianza relativamente alle cosiddette «trasformazioni di gauge» (per cui si veda ancora «Materia», pp. 921-22). Una opportuna teoria di gauge per i colori postula allora l'esistenza di otto particelle senza massa, dette gluoni (da *glue* 'colla'), che costituiscono i vettori della forza forte proprio come i fotoni lo sono per la forza elettromagnetica. Naturalmente, i gluoni, come i quark, non sono stati ancora «rivelati». Quando un quark emette o assorbe un gluone, cambia di colore ma non può cambiare di sapore.

I gluoni sono dunque i quanti per l'interazione forte (come i fotoni lo sono per l'interazione elettromagnetica): la teoria di gauge per i colori propone quindi che il vero carattere di tale interazione vada dunque cercato nella forza che lega insieme i quark colorati. Vale ora la pena di un confronto. La forza elettromagnetica tra due particelle cariche si esprime, notoriamente, con la legge di Coulomb, secondo cui la forza diminuisce come il quadrato della distanza tra le cariche. Una recente proposta sostiene che la forza forte tra due quark colorati si comporta in un modo differente: resta costante, indipendentemente dalla distanza tra i quark. In base a tale teoria, la quantità di energia necessaria a isolare un quark sarebbe enorme: inoltre, a partire dall'energia fornita per estrarre un quark isolato si materializzerebbero un nuovo quark e un nuovo antiquark. Il nuovo

quark rimpiazzerebbe allora il quark estratto, ricostruendo l'adrone originario, mentre il nuovo antiquark si combinerebbe con il quark estratto formando un mesone incolore.

In questa prospettiva si sarebbe tentati di asserire che la «cattiva infinità», il regresso all'infinito delle particelle – che si voleva fermare con l'ipotesi del *bootstrap*, soppiantata dal modello dei quark – potrebbe arrestarsi grazie a questa interpretazione del «confinamento» dei quark stessi. Gli atomi possono venir decomposti in elettroni e nuclei, i nuclei in protoni e neutroni, i protoni e i neutroni in quark. Ma a questo punto sembrerebbe difficile immaginare una struttura interna per delle «particelle» che non si riesce nemmeno a creare. La spettroscopia dei quark sarebbe dunque l'*ultima* spettroscopia, almeno per chi sottoscrive la teoria del confinamento di cui sopra: ma «l'esperienza del passato ci ha insegnato che affermazioni di questo genere si sono sempre rivelate inesatte» [Amaldi 1979, p. 117].

4.3. Quark e leptoni.

Nel 1976 il premio Nobel per la fisica è stato assegnato a Burton Richter (Stanford) e a Samuel Ting (Mit) per la scoperta (1974) di una particella dotata di proprietà piuttosto insolite, ottenuta da Ting e dal suo gruppo mediante collisioni di protoni di 28 GeV, prodotti dall'acceleratore di Brookhaven, con nuclei di berillio, e a Stanford con l'anello di accumulazione elettrone-positrone Spear, uno degli ultimi grandi discendenti di AdA. Battezzate J da Ting e ψ da Richter, tali particelle sono oggi note come particelle J/ψ. «Esse sembra inizino una nuova famiglia di particelle, diverse da tutte quelle conosciute prima perché, malgrado la loro considerevole massa (oltre tre volte quella del protone), sono relativamente stabili» [Segrè 1976, p. 256]. Va aggiunto subito che per le particelle J/ψ non sembra esservi posto nel modello a tre quark che si è finora esaminato. Ma già nel 1964, un anno dopo la proposta di Zweig e Gell-Mann, erano apparsi articoli che esponevano una teoria a quattro quark fondamentali. Dotata ancora di pochi agganci sperimentali, la teoria dei quattro quark conosce una rinascita solo nel 1970 (Glashow, Iliopoulos e Maiani) quando la simmetria $SU(4)$ sembra rivelarsi particolarmente promettente per l'interazione debole. Il quarto quark, battezzato in precedenza charm (c) da Glashow, spiega in modo estremamente semplice certe disintegrazioni deboli di particelle strane. c si suppone abbia carica $+2/3$ come u e inoltre porti un'unità $+1$ di un nuovo numero quantico (detto appunto charm), che si conserverebbe nella interazione forte e in quella elettromagnetica, ma non in quella debole. Ma esso si è imposto all'attenzione soprattutto identificando la particella J/ψ con la configurazione $c\bar{c}$ (si veda ancora «Materia», p. 926).

Qui interessa sottolineare come dopo le conferme sperimentali circa tale interpretazione di J/ψ «i due mondi distinti dei quark e dei leptoni paiono molto piú simmetrici» [Amaldi 1979, p. 119], come del resto risulta da un attento esame della tabella 2.

Nella tabella 2 compaiono anche il leptone τ e il suo neutrino ν_τ (Spear:

Stanford, 1975): giova ricordare che dal 1962 non si era prodotta alcuna aggiunta alla lista dei leptoni noti e la scoperta di un nuovo leptone con un'energia in riposo che raggiunge 1,8 GeV ha modificato notevolmente l'idea della elementarità di queste particelle [cfr. Perl 1978]. Nella tabella 2 compare anche un ulteriore quark, il quark b (un nuovo sapore) scoperto nel 1977 da Lederman e dal suo gruppo: basti qui dire che questo quinto quark si è manifestato, insieme con il suo antiquark, sotto la forma di un nuovo adrone (l'adrone *ipsilon*) dotato di lunga vita e di massa molto elevate (essendo dieci volte piú pesante del protone). Come lo stesso Lederman ha osservato [1978], tale scoperta può fare intravvedere qualche crepa nelle «fondamenta» dell'edificio delle particelle elementari cui si accennava a chiusura del § 4.2. Man mano che progredirà la tecnologia degli acceleratori, infatti, i fisici, con grande probabilità, troveranno sempre nuove entità subatomiche e gli stessi quark prolifereranno. Cinque, qualcuno dice già sei... e se ce ne sono sei perché non dodici, o ventiquattro, ecc.?

Ciò ha non pochi riflessi sulla trattatistica in cui gli stessi autori possono provare l'impressione di non riuscire a restare al passo di questo peculiare tipo di «progresso». «Mentre scrivevo questo libro, – avverte per esempio Close [1979, p. VI] in una prefazione datata 1978, – è stato scoperto un altro quark e si è verificata una vera e propria esplosione di idee». Si confronti del resto l'attuale situazione con quanto delineato nel § 3.4 a proposito della simmetria $SU(3)$. «Solo vent'anni fa la ricerca verteva sulla adeguata struttura gruppale per gli adroni. La moltitudine degli adroni che venne trovata aveva proprietà consistenti con l'esistenza di uno strato piú profondo della materia, quello contenente i quark, in origine proposti come *raison d'être* per la simmetria approssimata $SU(3)$ osservata nella spettroscopia degli adroni. Oggi crediamo che i quark e i

Tabella 2.

Quark, leptoni e mediatori delle forze fondamentali. Compaiono anche il quarto quark c (da *charm*) e il quinto quark b (da *beauty* 'bellezza'). È aperta la questione del sesto quark. Col termine *weakon* incluso tra i mediatori si intendono i bosoni intermedi (ai quali si è accennato nel § 3.1). La prima e la seconda famiglia comprendono tutti i quark e i leptoni scoperti prima del 1975 (Da Amaldi 1979).

		I famiglia	II famiglia	III famiglia
Particelle	Quark	u d	c s	? b
	Leptoni	ν_e e	ν_μ μ	ν_τ τ

		Forza fondamentale
Mediatori	*Weakon*	Interazione debole
	Fotoni	Interazione elettromagnetica
	Gluoni	Interazione forte

leptoni siano fondamentali e che le interazioni debole, elettromagnetica e forte possano venir descritte da un qualche gruppo che contenga $SU(2) \times U(1)$ e $SU(3)_{\text{colore}}$ come sottogruppi. Dunque ora la ricerca verte ancora una volta su quale deve essere la struttura di gruppo adeguata. Sembra che ci siano *almeno* cinque leptoni e quindici quark (cinque sapori per tre colori). Tutto ciò regge il paragone con il numero degli adroni noti quando $SU(3)$ venne per la prima volta proposto. Stiamo dunque entrando in una sorta di era della spettroscopia dei quark e dei leptoni che, alla fine, ci porterà oltre gli stessi quark?» [*ibid.*, p. 448]. [M. O'H.].

Amaldi, U.
1979 *Accélérateurs de particules et culture scientifique*, in V. Mathieu e P. Rossi (a cura di), *La culture scientifique dans le monde contemporain*, Scientia, Milano, pp. 43-141.

Bernardini, C.
1978 *Storia di AdA*, in «Scientia», CXIII, pp. 27-38.

Bloor, D. C.
1976 *Knowledge and Social Imagery*, Routledge and Kegan Paul, London.

Chen Ning Yang
1961 *Elementary Particles. A Short History of Some Discoveries in Atomic Physics*, Princeton University Press, Princeton N.J. (trad. it. Boringhieri, Torino 1969).

Chew, G. F.
1965 *Nuclear Physics: Democracy or Aristocracy*, in L. C. L. Yuan (a cura di), *Nature of Matter; Purposes of High Energy Physics*, Brookhaven National Laboratory, Upton N.Y., pp. 35-40.

Close, F. E.
1979 *An Introduction to Quarks and Particles*, Academic Press, New York.

Dirac, P.-A.-M.
1974 «Dirac, Paul Adrien Maurice», in *Scienziati e tecnologi contemporanei*, vol. I, Mondadori, Milano, pp. 310-12.

Dyson, F. J.
1964 *Mathematics and Physical Sciences*, in «Scientific American», CCXI, 3, pp. 128-46; ora in *The Mathematical Sciences; a Collection of Essays*, Mit Press, Cambridge Mass. 1969 (trad. it. Zanichelli, Bologna 1973, pp. 114-34).

Einstein, A.
1905 *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt*, in «Annalen der Physik», IV, 17, pp. 132-48 (trad. it. in *La teoria dei quanti di luce*, Newton Compton, Roma 1972, pp. 45-72).

Flower, W. B., e Samios, N. P.
1964 *The Omega-minus experiment*, in «Scientific American», CCXI, 4, pp. 36-45.

Ford, K. W.
1963 *The World of Elementary Particles*, Blaisdell, New York (trad. it. Mondadori, Milano 1975[4]).

Hanson, N. R.
1958 *Patterns of Discovery. An Inquiry into the Conceptual Foundations of Science*, Cambridge University Press, London (trad. it. Feltrinelli, Milano 1978).
1963 *The Concept of the Positron: a Philosophical Analysis*, Cambridge University Press, London.

Kuhn, Th. S.
1962 *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago (trad. it. Einaudi, Torino 1976[3]).

Lakatos, I.
1970 *Falsification and the Methodology of Scientific Research Programmes*, in I. Lakatos e A. M. Musgrave (a cura di), *Criticism and the Growth of Knowledge*, Cambridge University Press, London, pp. 91-196 (trad. it. Feltrinelli, Milano 1976, pp. 164-276).

Lederman, L. M.
1978 *The upsilon particle*, in «Scientific American», CCXXXIX, 4, pp. 60-68 (trad. it. in «Le Scienze», XI (1978), 124, pp. 22-31).

Matthews, P. T.
1971 *The Nuclear Apple: Recent Discoveries in Fundamental Physics*, Chatto and Windus, London (trad. it. Mondadori, Milano 1976[2]).

Neumann, J. von
1932 *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin.

Omnès, R.
1970 *Introduction à l'étude des particules élémentaires*, Ediscience, Paris.

Oppenheimer, J. R.
1954 *Science and the Common Understanding*, Simon and Schuster, New York (trad. it. Boringhieri, Torino 1965).

Pearson, K.
1911 *The Grammar of Science*, Black, London.

Perl, M.
1978 *The tau heavy lepton: a recently discovered elementary particle*, in «Nature», CCLXXV, pp. 273-78.

Segrè, E.
1976 *Personaggi e scoperte nella fisica contemporanea*, Mondadori, Milano.

Shapere, D.
1969 *Notes toward a Post-Positivistic Interpretation of Science*, in P. Achinstein e S. F. Barker (a cura di), *The Legacy of Logical Positivism; Studies in the Philosophy of Science*, Johns Hopkins Press, Baltimore, pp. 115-60.

Toraldo di Francia, G.
1976 *L'indagine del mondo fisico*, Einaudi, Torino.

Vaihinger, H.
1911 *Die Philosophie des 'Als Ob'*, Meiner, Leipzig (trad. it. Ubaldini, Roma 1967).

Weyl, H.
1952 *Simmetry*, Princeton University Press, Princeton N.J. (trad. it. Feltrinelli, Milano 1962).

Yukawa, H.
1974 «Yukawa, Hideki», in *Scienziati e tecnologi contemporanei*, vol. III, Mondadori, Milano, p. 198.

Dall'atomismo democriteo (cfr. **atomo e molecola**) alla **fisica** contemporanea, l'individuazione delle particelle elementari (cfr. anche **elementi**), intese come costituenti ultimi della **materia** (cfr. anche **luce**, **gravitazione**) si è imbattuta non solo in difficoltà tipicamente teoriche (cfr. **teoria/modello**) ma ha costituito una sfida per le stesse capacità dell'**osservazione** (cfr. **dato**, **esperimento** e anche **empiria/esperienza**). Il quadro concettuale (cfr. **paradigma**) fornito dalla **relatività** e dalla meccanica dei **quanti** ha infine consentito, nel Novecento, un'ampia modellistica (cfr. **modello**) dei fenomeni microfisici grazie sia all'impiego di **matematiche** notevolmente sofisticate (cfr. **calcolo**, **numero** e anche **applicazioni**, **funzioni** e **strutture matematiche**) sia all'utilizzazione di un apparato sperimentale estremamente potente (cfr. **macchina**,

tecnica). Entro il **vincolo** delle leggi (cfr. **legge**) di conservazione (cfr. **conservazione/invarianza**) e attraverso lo studio della **simmetria** fisica (cfr. anche **invariante**) è emersa una **sistematica e classificazione** dello «zoo delle particelle» stabili e instabili (cfr. **stabilità/instabilità**), mentre la ricerca si muove oggi nella direzione di una sintesi in grado di cogliere l'**unità** profonda delle differenti forme fondamentali di **interazione** (cfr. anche **forza/campo**). Questa concezione unitaria (cfr. anche **uno/molti**, e per altri aspetti **identità/differenza**) della realtà fisica (cfr. anche **reale**) è oggi però piú un **programma** che un risultato della **ricerca**.

Plasma

In questo articolo la parola 'plasma' indica il quarto stato della materia, in cui un gas viene trasformato in condizioni di elevata temperatura e bassa densità. La sua etimologia (dal greco πλάσσειν 'plasmare') e le sue risonanze semantiche sono purtroppo fuorvianti: esse si ricollegano alle primissime fasi della ricerca in questo campo quando, studiando le scariche elettriche nei gas (come quelle che hanno luogo negli ordinari tubi al neon), si osservò che esse dànno origine a zone di diffusa luminosità e vengono modificate in forma e in splendore da campi magnetici ed elettrici. Queste ricerche ebbero inizio nel secolo scorso e furono soprattutto coltivate dalla scuola sperimentale inglese, in particolare da Faraday; e si accumulò un grande e articolato patrimonio semiempirico sui complessi fenomeni che hanno luogo quando un gas è attraversato da una corrente elettrica. Da allora questa disciplina fisica è straordinariamente cresciuta nell'elaborazione concettuale e nella strumentazione di laboratorio; ai nostri occhi le ricerche di fisica del plasma fino a prima della seconda guerra mondiale hanno solo interesse storico. Il nome 'plasma' è oggi giustificato solo dall'uso corrente e non dalla sua etimologia e sta ad indicare un gas ionizzato (parzialmente o completamente). La ionizzazione di un atomo ha luogo quando uno o piú degli elettroni che ruotano attorno al nucleo vengono strappati; si producono cosí cariche elettriche libere che modificano profondamente le proprietà della materia. In un atomo ordinario vi è esatta neutralità di carica: nel nucleo sono presenti tanti protoni (con carica e positiva) quanti sono gli elettroni orbitali (ciascuno con carica $-e$). Un gas ionizzato è ancora in media neutro, ma si possono verificare deviazioni dalla neutralità anche su regioni assai estese; e, soprattutto, esso risente fortemente delle forze elettriche e magnetiche esterne.

Il fenomeno della ionizzazione, messo in evidenza nella seconda metà del secolo scorso, ha segnato il primo passo nella falsificazione dell'atomismo che aveva dominato la chimica nell'età precedente: gli atomi, che si comportano come unità indivisibili nelle reazioni chimiche, sono in realtà formati da parti piú piccole. Questo processo di disaggregazione dei componenti della materia è poi proseguito a livelli inferiori, prima con la scomposizione del nucleo nei suoi costituenti – i nucleoni, che sono di due specie, i protoni e i neutroni – e poi con la scoperta o l'ipotesi di particelle subnucleari (come i cosiddetti *quark*, tre dei quali sono necessari per costituire un nucleone). Ciascuna di queste unità composte è caratterizzata da una forza che ne tiene insieme i membri e non si manifesta normalmente all'esterno: cosí le forze interatomiche legano gli atomi nelle molecole, le forze elettriche agiscono tra i costituenti dell'atomo e le interazioni forti tra i nucleoni nel nucleo. Ognuno di questi sistemi è caratterizzato dall'energia necessaria per separarne i membri; per esempio, nell'atomo d'idrogeno è necessaria l'energia di 13,6 eV per strappare al nucleo l'unico elettrone presente. Un elettronvolt (abbreviato eV) è l'energia che un elettrone acquista quando viene accelerato in un potenziale elettrico di un volt. Per tradurre questa unità in

termini piú accessibili, si osservi che l'energia cinetica di agitazione termica di 1 eV corrisponde alla temperatura di 11 000 gradi; quindi a 150 000 gradi ogni atomo d'idrogeno può essere ionizzato da una collisione. Per la maggior parte degli altri elementi l'energia di ionizzazione è inferiore, ma queste cifre mostrano subito che non è facile ottenere materia ionizzata in condizioni ordinarie.

La crescita tumultuosa e talora costosissima della fisica del plasma negli ultimi decenni è dovuta principalmente a due motivi, uno economico e uno scientifico. Come si vedrà nel § 4, per produrre in laboratorio le reazioni di fusione che hanno luogo nel Sole e produrre cosí energia nucleare dall'acqua del mare occorre ottenere temperature altissime, alle quali la materia è certamente ionizzata. Questo progetto, che potrà rendere possibile nei prossimi decenni una nuova ed inesauribile fonte di energia, ha però incontrato ed incontra tuttora difficoltà di carattere fisico, che possono venire superate solo studiando a fondo la fisica del plasma ad alta temperatura. Sono state cosí messe a disposizione notevolissime risorse finanziarie, umane e sperimentali che hanno cambiato qualitativamente l'aspetto di questa disciplina.

Il motivo scientifico (cfr. § 3) sta nell'importanza che il plasma ha nei fenomeni cosmici e nella grande varietà di effetti a cui dà luogo. Da quando negli anni '30 è apparso chiaro che l'energia radiante delle stelle e, in particolare, del Sole è dovuta alle reazioni termonucleari nel loro interno, si è riconosciuto che il plasma è il principale costituente dell'universo e la forma normale della materia.

La fisica del plasma ha cosí raggiunto sul piano concettuale una sua compiuta elaborazione; essa si fonda sulla meccanica classica, l'elettromagnetismo e la teoria quantistica, ma costituisce una disciplina a sé stante, con i suoi paradigmi teorici, i suoi metodi di calcolo e di analisi sperimentale. In questa ricerca non vi sono incertezze di base, non vi sono leggi fondamentali da scoprire; ma la straordinaria varietà di effetti, di situazioni e di fenomeni, originata anche dalle applicazioni astrofisiche ed energetiche, ne hanno fatto un campo vivacissimo in rapido e proliferante sviluppo. Nella terminologia usata da Kuhn nella sua *Structure of Scientific Revolutions* (1962), la fisica del plasma è una «scienza normale», che si sviluppa mediante una sistematica accumulazione di risultati secondo precise prassi ed esempi codificati nei libri di testo ed accettati senza discussione dalla comunità dei ricercatori. In essa non vi sono anomalie e contraddizioni di fondo, che ne possono compromettere la stabilità e la crescita progressiva.

1. *Proprietà del plasma.*

In questo paragrafo saranno delineate quattro proprietà caratteristiche di questo stato della materia. Esse, come si vedrà, dànno luogo a chiare rappresentazioni intuitive e rendono assai piú facile la comprensione di fenomeni complessi e rapide stime degli ordini di grandezza; si può dire che nel lavoro del ricercatore professionista in questo campo queste proprietà sono familiari strumenti di lavoro a cui egli ricorre quasi inconsciamente, prima ancora di arrivare alla formulazione matematica precisa.

1.1. Neutralità di carica.

Il plasma è in media elettricamente neutro perché per separare le cariche positive da quelle negative occorre spendere energia. L'energia cinetica media di una particella di gas alla temperatura T è $3/2\,kT$ (k è la costante di Boltzmann); essa, per il teorema di equipartizione dell'energia, sarà anche circa uguale all'energia potenziale elettrostatica media provocata da una carenza o da una sovrabbondanza di cariche di un certo segno attorno alla particella presa in considerazione. Se il plasma contiene n elettroni per cm^3, un vuoto di elettroni di dimensioni d ha una carica positiva dell'ordine end^3 e, quindi, conferisce a un elettrone un'energia potenziale $e^2nd^3/d = e^2nd^2$. Uguagliando questa espressione all'energia termica kT, si ottiene la dimensione media dei vuoti di carica:

$$d = \sqrt{\frac{kT}{e^2n}}.$$

Questa lunghezza caratteristica prende il nome da Debye (nella terminologia corrente la lunghezza di Debye è $d/\sqrt{4\pi}$) ed è, nei casi ordinari dei plasmi di laboratorio, assai piccola. Ovviamente questo ragionamento vale solo se in un volume di dimensioni d vi sono molte particelle, cioè se $nd^3 \gg 1$. Il termine 'plasma', nella sua accezione ordinaria, è riservato ai sistemi in cui questa disuguaglianza è verificata.

Questa considerazione mostra anche che, in generale, le particelle di un plasma non interagiscono individualmente (a meno che esse non subiscano una vera e propria collisione, il che accade raramente); piuttosto, le fluttuazioni di carica di dimensioni d che si formano e si disfano continuamente si attraggono e si respingono tra di loro anche a grande distanza. In altre parole, il plasma mostra un comportamento *collettivo* in cui gli effetti macroscopici sono causati dal movimento coordinato di un grande numero di particelle. L'analisi di questi fenomeni può essere fatta solo con raffinati metodi statistici.

Allo scopo di dare maggior concretezza alle formule algebriche, se ne darà il valore numerico per un tipico plasma di laboratorio, con $n = 10^{14}$ elettroni per cm^3, $T = 10^6$K (K indica i gradi Kelvin, misurati a partire dallo zero assoluto) e nuclei costituiti da un solo protone (idrogeno ionizzato). Allora $d/\sqrt{4\pi} = 6 \times 10^{-3}$ cm.

Le fluttuazioni di una carica possono essere prodotte anche artificialmente, per esempio con un elettrodo il cui potenziale oscilla a una frequenza ω. Si trova che il plasma possiede una frequenza caratteristica

$$\omega_p = \sqrt{\frac{4\pi e^2 n}{m_e}}$$

($= 5 \times 10^{11}$ sec^{-1}; m_e è la massa dell'elettrone) alla quale esso può risuonare, indipendentemente dalle sue dimensioni (a differenza dei sistemi materiali solidi, che dànno una nota tanto piú acuta quanto piú sono piccoli). Se la frequenza eccitatrice è inferiore a ω_p la perturbazione viene subito smorzata; se è assai supe-

riore la sua velocità di propagazione V è sostanzialmente la velocità termica degli elettroni $\sqrt{kT/m_e}$ ($=4\times10^8$ cm sec^{-1}), ma diventa sempre piú grande man mano che ci s'avvicina a ω_p (fig. 1):

$$V=\sqrt{\frac{3kT}{m_e}\frac{1}{1-\frac{\omega_p^2}{\omega^2}}}.$$

Queste oscillazioni di plasma sono il capostipite di una numerosa e complessa famiglia di moti ondosi prodotti dal moto collettivo delle particelle nei campi elettrici e magnetici generati da esse stesse ed eventualmente prodotti dall'esterno. Un altro esempio fondamentale si ha quando un'onda elettromagnetica di frequenza e lunghezza d'onda $\lambda=2\pi c/\omega$ incide su un plasma: al suo interno la lunghezza d'onda

$$\lambda=\frac{2\pi c}{\sqrt{\omega^2-\omega_p^2}}$$

diviene sempre piú grande fino a diventare formalmente infinita per $\omega=\omega_p$. Per frequenze inferiori λ diventa immaginaria: il che significa che l'onda è smorzata e non è in grado di penetrare il plasma.

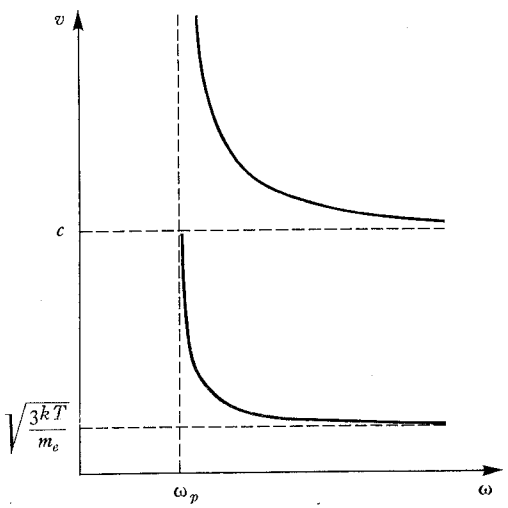


Figura 1.

La velocità di propagazione v di due caratteristici fenomeni ondulatori in un plasma. La curva superiore corrisponde alle onde elettromagnetiche, che nel vuoto si propagano con la velocità della luce c; quella inferiore vale nel caso di oscillazioni della carica elettrica del mezzo. Ambedue le onde non possono propagarsi se la loro frequenza ω è inferiore alla frequenza di plasma

$$\omega_p=\left(\frac{4\pi ne^2}{m_e}\right)^{1/2}.$$

1.2. Magnetismo.

Nei plasmi ordinari il campo magnetico prodotto dalle singole particelle è trascurabile; un campo magnetico esterno, invece, è in grado di esercitare effetti drastici.

Esso curva la traiettoria di solito rettilinea di una carica in un'elica (fig. 2) il cui asse è allineato con le linee di forza; la velocità angolare di rotazione è uguale alla cosiddetta «frequenza di ciclotrone» $\Omega=eB/mc$ (si usano qui unità elettrostatiche assolute per la carica e il gauss per il campo magnetico). Per un elettrone in un campo magnetico di 10 000 gauss, $\Omega=1{,}7\times10^{11}$ sec^{-1}, un po' inferiore alla frequenza di plasma. Il raggio dell'elica è il rapporto tra la velocità $v_\perp$ ortogonale a **B** e Ω e può essere piccolo; per un protone a 10^6K esso è 1 cm. Si può dire che le particelle vengono, per cosí dire, ancorate alle linee di forza, pur essendo sempre libere di muoversi lungo di esse. I tre gradi di libertà traslazionali delle molecole di un gas ordinario (o di un plasma non magnetizzato) sono qui ridotti a uno solo.

Quando un flusso di plasma viene diretto verso un campo magnetico ortogonale alla sua velocità il suo moto viene ostacolato; se il flusso è intenso esso gene-

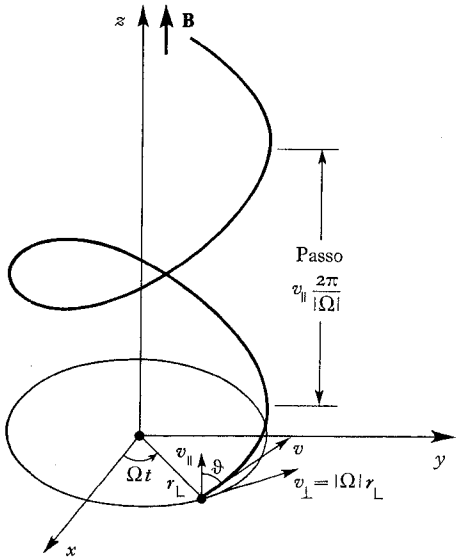


Figura 2.

Traiettoria di una particella carica in un campo magnetico uniforme e costante. L'elica è descritta con un moto di traslazione uniforme lungo l'asse, parallelo al campo con velocità $v_\parallel$; e con un moto circolare uniforme nel piano perpendicolare con frequenza $\Omega=eB/mc$ e velocità $v_\perp$. L'angolo ϑ è formato tra la velocità complessiva e **B**. Il raggio dell'elica $r_\perp=v_\perp/|\Omega|$ si dice «raggio di Larmor».

ra un campo magnetico di segno opposto a quello esterno, diminuendolo o addirittura annullandolo. Lo studio di questa competizione è uno dei problemi principali di una parte della fisica del plasma, la «magnetofluidodinamica», generalizzazione della dinamica dei fluidi ordinari. Esso è correntemente risolto attraverso il concetto di *pressione magnetica*: un plasma è soggetto, oltre alla sua pressione idrostatica P, a una pressione prodotta dal campo magnetico esterno **B** data da

$$P_m = \frac{B^2}{4\pi},$$

che agisce nel piano perpendicolare a **B**. Per mantenere un plasma in equilibrio occorre assicurare che la pressione totale $P+P_m$ sia costante; si può allora impedire al plasma di fluire attraverso linee di forza circondandolo con un campo magnetico esterno. Su questo principio si fondano molti dispositivi sperimentali per confinare e riscaldare un plasma; difatti non è opportuno che esso venga a contatto con una parete solida perché verrebbe da essa rapidamente raffreddato e contaminato da impurità estratte dai primi strati superficiali.

Ogni configurazione di confinamento magnetico deve però risolvere il problema del moto delle particelle lungo le linee di forza. Anzitutto si può creare un campo magnetico di forma anulare senza estremità (i cosiddetti «tori»: cfr. fig. 3); ma si può anche sfruttare un'importante proprietà del moto di una carica in un campo magnetico non uniforme. Quando B varia l'elica descritta è lievemente deformata, ma rimane costante il rapporto $B/\sin^2\vartheta$, ove ϑ è l'angolo che la velocità della particella fa con l'asse (fig. 2). All'aumentare di B, quindi, il passo dell'elica diventa sempre piú piccolo e ad un certo punto comincia ad avvolgersi all'indietro, invertendo il senso di marcia lungo la linea di forza. Le zone ove B

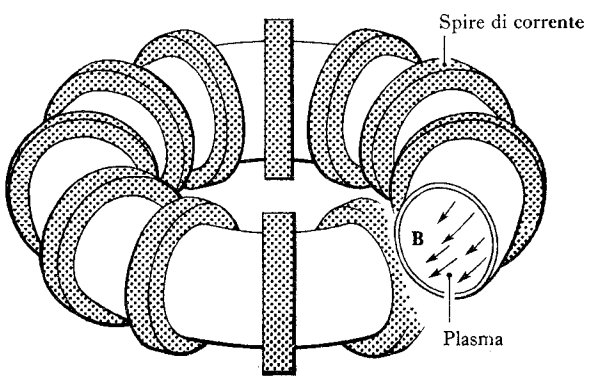


Figura 3.

Una maniera per eliminare perdite di particelle dalle estremità delle linee di forza magnetiche: esse vengono chiuse ad anello e il plasma racchiuso – senza toccare le pareti – in un contenitore a forma di «toro». Le spire esterne generano un campo magnetico allineato con l'asse del toro.

è piú intenso si comportano quindi come degli specchi che riflettono all'indietro gran parte delle particelle. Su questo principio si basa un altro importante sistema di confinamento del plasma, la bottiglia magnetica (fig. 4).

1.3. Trasporto.

Come è naturale, data la sua composizione, un plasma è un ottimo conduttore di corrente elettrica. Un campo elettrico esterno **E** (per esempio, quello creato da un condensatore) imprime a ogni carica la forza $e\mathbf{E}$; essa si muove quindi di moto accelerato fino al momento in cui una collisione con un'altra particella le conferisce una velocità completamente casuale e, in media, nulla. Da questo istante in poi riprende l'accelerazione, e cosí via. Indicando con τ l'intervallo di tempo che in media separa due collisioni successive, la velocità media ordinata acquistata dalla particella è

$$v = \frac{1}{2}\frac{eE}{m}\tau;$$

attraverso una sezione di area unitaria passerà allora una corrente

$$j = env = \frac{1}{2}\frac{e^2E\tau n}{m}$$

(equivalente alla legge di Ohm).

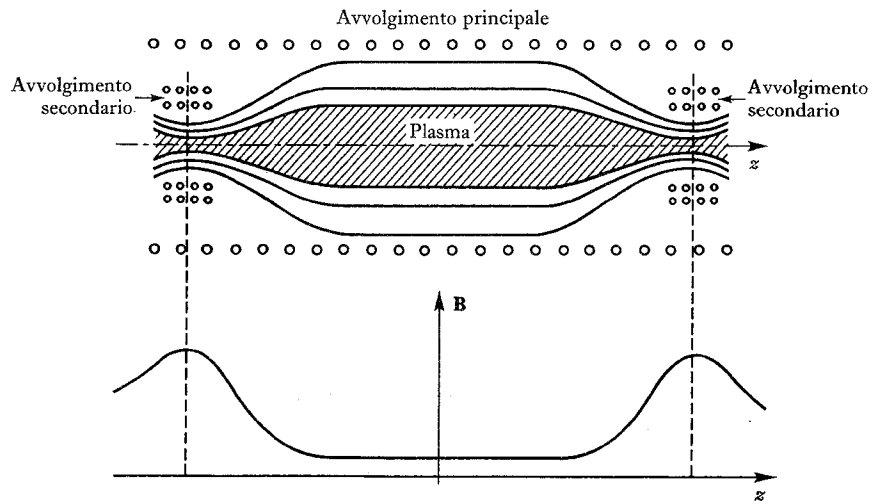


Figura 4.

Un altro metodo per ottenere un plasma confinato: la «bottiglia magnetica». L'avvolgimento principale produce un campo con simmetria cilindrica; i due gruppi di spire secondarie lo rendono piú intenso alle estremità del plasma. I due grafici indicano le linee di forza e l'andamento dell'intensità del campo sull'asse.

Si vede che questa corrente è proporzionale al campo elettrico esterno; e il coefficiente di proporzionalità, che si dice «conducibilità elettrica» è, $\sigma = e^2\tau n/2m$. La valutazione del tempo τ costituisce uno dei capitoli piú delicati di tutta la fisica del plasma; basti qui dire che una rozza approssimazione è data dalla formula $\tau\omega_p \approx n(d/\sqrt{4\pi})^3$. Al secondo membro si ha il numero di elettroni contenuti in un cubo il cui spigolo è uguale alla lunghezza di Debye; esso, come si è detto, è di solito assai grande (nel caso specifico sopra indicato è circa 100 000). Si vede allora che tra una collisione e l'altra trascorrono moltissimi periodi delle oscillazioni di plasma, il che rende ragione del permanere della loro coerenza per tempi lunghi.

Per ragioni simili il plasma è un ottimo conduttore di calore; anche in questo caso svolge un ruolo importante la distanza che in media una particella copre prima di cedere in una collisione l'energia termica che possiede alle altre cariche del gas. Tale distanza è uguale al prodotto di τ per la velocità media ed è quindi assai grande quando vi sono molte particelle in un cubo di Debye (di spigolo d).

1.4. Irraggiamento.

L'irraggiamento di un plasma è assai diverso nei due casi in cui esso sia opaco o trasparente alla radiazione generata. Se è opaco ogni fotone collide ben presto con un'altra carica e si pone in una situazione di equilibrio termodinamico in cui la sua energia media è uguale a kT, ove T è la temperatura del plasma. Allora l'emissione radiativa è proporzionale alla superficie del plasma e il flusso emesso per unità di superficie è proporzionale a T^4. Questo è il caso del Sole, il quale non è altro che una grande massa sferica di plasma d'idrogeno ed elio opaco alla sua stessa radiazione.

Se invece si hanno condizioni di trasparenza, ogni fotone emesso è perso; la potenza della radiazione è proporzionale al volume del plasma ed è determinata dai meccanismi dettagliati con cui l'energia delle singole particelle si trasforma in energia elettromagnetica. Si accenna qui a due di questi meccanismi, basati su un importante risultato dell'elettrodinamica classica: ogni particella in moto accelerato irraggia, e la potenza radiativa è proporzionale al quadrato dell'accelerazione. In un plasma vi sono due cause principali dell'accelerazione di una carica. Anzitutto quando, ad esempio, un elettrone passa in prossimità di un nucleo, viene deviato dal suo campo elettrostatico e quindi irraggia per un periodo che è dell'ordine della sua distanza minima divisa per la velocità. In conseguenza di questa radiazione l'energia dell'elettrone diminuisce ed esso viene frenato; essa si chiama pertanto «radiazione di frenamento». La potenza emessa per centimetro cubo è proporzionale a $n^2\sqrt{T}$ e diventa rapidamente importante all'aumentare della densità.

Se sul plasma agisce un campo magnetico esterno ogni particella ha un'accelerazione $v_\perp\Omega$ e la potenza corrispondente è proporzionale a B^2. Essa si chiama «radiazione di ciclotrone» ed è emessa proprio alla frequenza Ω. Quando la velocità delle particelle è prossima a quella della luce, tuttavia, questo tipo di radiazione acquista delle caratteristiche assai peculiari a motivo degli effetti relativistici; per esempio, la radiazione è polarizzata e diretta lungo raggi che formano con la traiettoria un piccolo angolo. Essa si chiama «radiazione di sincrotrone». Spesso i plasmi che si incontrano in astrofisica sono prossimi a intense sorgenti di particelle assai energetiche, come le stelle di neutroni, i nuclei galattici o il Sole stesso; la radiazione che si osserva in questi casi ha una componente di sincrotrone e la sua analisi costituisce uno strumento prezioso per risalire alle proprietà della sorgente.

2. *Livelli di descrizione matematica del plasma.*

Dopo aver passato in breve rassegna le proprietà e le raffigurazioni intuitive che guidano la comprensione della fisica del plasma, è interessante ora analizzare i diversi livelli di descrizione matematica che generano i potenti metodi di calcolo oggi usati per predire quantitativamente il suo comportamento. Poiché le leggi di moto per ciascuna carica sono perfettamente note, in linea di principio si può ottenere del plasma una descrizione esatta e deterministica, capace di dare le posizioni e le velocità di tutte le particelle – e quindi anche tutte le proprietà macroscopiche – quando siano note le posizioni e le velocità all'istante iniziale. In realtà ciò non è possibile nei casi realistici per l'elevatissimo numero dei gradi di libertà che intervengono; occorre adottare precise approssimazioni, la cui validità è tuttavia soggetta, talvolta, a dubbi. Negli ultimi anni hanno assunto grande importanza le simulazioni numeriche della dinamica del plasma, in cui con grandi elaboratori elettronici si sono risolte le equazioni di moto di un insieme di particelle; il loro numero (per esempio, qualche migliaio) è sempre di gran lunga inferiore ai valori realistici, ma i risultati costituiscono un utile controllo delle conclusioni teoriche e qualche volta hanno indicato nuovi effetti.

Sia dal punto di vista dell'evoluzione storica sia da quello della naturale progressione della difficoltà, il primo modello di plasma è quello di un fluido in grado di trasportare una corrente elettrica e soggetto all'azione delle forze esterne e dei campi magnetici da esso stesso generati. Per descriverlo occorre specificare per ogni punto e per ogni istante le seguenti grandezze: la densità di massa, la velocità di flusso, la temperatura (o meglio, l'energia cinetica media per particella) e il campo magnetico. Ciascuna di queste grandezze obbedisce a una legge fisica che permette di calcolarne la variazione in un piccolo intervallo di tempo. Piú precisamente, la densità di massa è determinata dalla conservazione della massa contenuta in ogni volumetto di fluido; la velocità di flusso deve soddisfare la conservazione della quantità di moto; la temperatura è governata dalla conservazione dell'energia tenendo conto, naturalmente, di tutti i processi di trasporto; e, infine, il campo magnetico è legato dalla legge di induzione al campo elettrico, a sua volta determinato dalla corrente trasportata con la legge di Ohm già ricordata. Non è necessario per comprendere lo schema scrivere in dettaglio queste equazioni; basta ricordare che esse determinano l'evoluzione temporale di due grandezze vettoriali (la velocità di flusso e il campo magnetico) e di due grandezze scalari (la densità di massa e la temperatura). Ad esse si deve aggiun-

gere la definizione della pressione e della conduttività in funzione delle altre grandezze.

Questo schema, che va sotto il nome piuttosto ostico e già citato di magnetofluidodinamica, si è dimostrato straordinariamente fecondo ed ha permesso, per esempio, di chiarire il meccanismo che dà origine al campo magnetico terrestre, come si accennerà nel § 3.

Ad un'analisi piú approfondita, tuttavia, le limitazioni del modello fluido del plasma appaiono ben presto chiare. Anzitutto esso considera un fluido elettricamente neutro e non è quindi in grado di descrivere le fluttuazioni di carica e le oscillazioni di plasma. Si è visto che le prime hanno dimensioni spaziali dell'ordine della lunghezza di Debye; e le seconde hanno frequenze superiori alla frequenza di plasma. Pertanto questo modello è sensato solo per fenomeni che avvengono su una scala di tempo piú lunga del reciproco di ω_p; e su dimensioni spaziali piú grandi di d, sulle quali le deviazioni dalla neutralità sono trascurabili. In secondo luogo, nell'ambito di questo modello non possono venire descritti il processo di ionizzazione e il suo inverso di ricombinazione, che possono mettere in gioco notevoli quantità di energie. Essi richiedono la considerazione di due fluidi distinti e sovrapposti, uno neutro e un altro ionizzato, con la possibilità di scambi di materia tra di loro. Ma esso ha anche una carenza assai piú sottile, la cui comprensione introduce nel cuore della fisica del plasma.

Occorre domandarsi: in quali condizioni la descrizione del plasma come un fluido può dirsi incompleta? Quando si asserisce che in un punto di un gas ordinario (ove, per semplicità si suppone nulla la velocità del flusso) vi è una temperatura T, non si dice solamente che l'energia cinetica media delle molecole che lo compongono è $3/2\,kT$, ma si fa in realtà un'asserzione assai piú complessa circa la loro distribuzione statistica. Precisamente, si richiede *a*) che le direzioni della loro velocità siano tutte ugualmente probabili (isotropia) e *b*) che la probabilità che la loro energia cinetica sia compresa tra E e $E+dE$ sia data esattamente dalla distribuzione di Boltzmann, cioè da

$$\sqrt{\frac{4E}{\pi}}\,(kT)^{-3/2}e^{-E/kT}\,dE$$

(cfr. fig. 5). In questa maniera l'assegnazione della temperatura e della densità determina completamente lo stato statistico del gas in quel punto. Questa distribuzione di velocità è privilegiata perché è il risultato naturale dell'azione di disordinamento prodotta nelle particelle dalle collisioni; in altre parole, se in partenza esse hanno una distribuzione diversa (ad esempio, non isotropa), dopo un tempo dell'ordine di τ (durante il quale la maggior parte delle particelle subiscono almeno una collisione) essa viene normalizzata ed assume le due caratteristiche *a*) e *b*) indicate. Si è qui trascurata, per semplicità, la distinzione tra il rilassamento alla proprietà *a*) e quello alla proprietà *b*). Quando, come accade di solito, le particelle negative hanno massa diversa da quelle positive, essi sono caratterizzati da tempi diversi. Questo stato finale si chiama equilibrio termodinamico locale e costituisce, per cosí dire, lo stato di massimo disordine locale possibile; e il processo che ad esso conduce è il «rilassamento verso l'equilibrio».

Allora se i fenomeni che si vogliono descrivere avvengono su una scala di tempi piú grande del tempo τ, ogni eventuale deviazione dall'isotropia e dalla distribuzione di Boltzmann ha il tempo di venire cancellata dalle collisioni che si susseguono numerosissime e caotiche.

Nel plasma tipico di laboratorio considerato, τ è dell'ordine di 3×10^{-5} sec, un tempo certo piccolo per i fenomeni ordinari, ma assai piú lungo del periodo di plasma $1/\omega_p$ e del periodo di ciclotrone $1/\Omega$. Quindi in questo caso tale approssimazione non è adatta a descrivere le oscillazioni di plasma e tutti gli altri fenomeni di frequenza elevata. Per essi vale piuttosto l'approssimazione opposta, in cui le collisioni vengono completamente trascurate; ma allora occorre adottare una descrizione assai piú dettagliata dello stato cinematico delle particelle, che ammetta un arbitrario grado di anisotropia e di deviazione dalla distribuzione di Boltzmann. Tale descrizione richiede l'introduzione, oltre allo spazio ordinario (di coordinate cartesiane x,y,z), dello spazio le cui coordinate sono le componenti della velocità v_x, v_y, v_z; e di una funzione di densità in questo spazio complessivo a sei dimensioni. Un plasma è allora descritto da un «fluido» continuo che si muove in questo spazio. Matematicamente questa funzione densità si ottiene considerando un volumetto $dx\,dy\,dz$ nello spazio ordinario attorno al punto (x,y,z) e un volumetto $dv_x\,dv_y\,dv_z$ nello spazio delle velocità attorno al punto (v_x,v_y,v_z). Se dN è il numero totale delle particelle presenti in questo volumetto a sei dimensioni, la densità f (funzione di distribuzione) è il rapporto

$$\frac{dN}{dx\,dy\,dz\,dv_x\,dv_y\,dv_z}=f(x,y,z,v_x,v_y,v_z)$$

quando il denominatore tende a zero. In realtà, naturalmente, il plasma è costituito da particelle singole e questo passaggio al limite non è a rigor di termini possibile; si può allora definire la funzione di distribuzione quando il volumetto è abbastanza grande da contenere un gran numero di particelle e abbastanza piccolo affinché in esso non vi siano apprezzabili variazioni delle quantità macroscopiche. Con un procedimento piú rigoroso, ma astratto, si possono anche immaginare infinite repliche ideali dello stesso sistema fisico e intendere f come la densità di probabilità nello «spazio delle fasi» (x, y, z, v_x, v_y, v_z).

L'evoluzione temporale della f è determinata da un'equazione (detta di Vlasov dal nome del fisico russo che l'ha dedotta) che si deduce dalla dinamica delle singole particelle. In un tempo infinitesimo dt una carica che si trova in (x,y,z,v_x,v_y,v_z) si sposta in $(x+v_x dt,\ y+v_y dt,\ z+v_z dt,\ v_x+(F_x/m)\,dt,\ v_y+(F_y/m)\,dt,\ v_z+(F_z/m)\,dt)$, ove $\mathbf{F}=(F_x,F_y,F_z)$ è la forza agente; ciò produce una variazione della f descritta dall'equazione

$$\frac{\partial f}{\partial t}+v_x\frac{\partial f}{\partial x}+v_y\frac{\partial f}{\partial y}+v_z\frac{\partial f}{\partial z}+\frac{F_x}{m}\frac{\partial f}{\partial v_x}+\frac{F_y}{m}\frac{\partial f}{\partial v_y}+\frac{F_z}{m}\frac{\partial f}{\partial v_z}=0.$$

Il punto cruciale però sta nella maniera con cui si calcola la forza $\mathbf{F}$: essa è la somma delle forze esterne e di quelle dovute al plasma stesso, dipendenti quindi dalla funzione di distribuzione f. L'equazione che cosí si ottiene presenta caratteristici effetti di azione e reazione: le forze che modificano la f sono a loro volta

determinate dalla f attraverso la cariche e le correnti che il plasma contiene. Se f_1 e f_2 sono due soluzioni dell'equazione di Vlasov, la loro somma f_1+f_2 in generale non la soddisfa: essa è, come si dice, non-lineare.

Nota la funzione di distribuzione (ad esempio, degli elettroni, con simbolo f_e), è immediato il calcolo delle grandezze fluide corrispondenti. La densità – cioè il numero delle particelle per unità di volume – è

$$n_e=\int_{-\infty}^{+\infty} dv_x \int_{-\infty}^{+\infty} dv_y \int_{-\infty}^{+\infty} dv_z\, f_e;$$

la velocità media $\mathbf{u}_e$ è proporzionale alla quantità di moto per unità di volume

$$n_e\mathbf{u}_e=\int_{-\infty}^{+\infty} dv_x \int_{-\infty}^{+\infty} dv_y \int_{-\infty}^{+\infty} dv_z\, \mathbf{v} f_e;$$

l'energia media T_e (che con una generalizzazione del termine si può chiamare temperatura) è data da

$$\frac{3}{2} n_e k T_e=\int_{-\infty}^{+\infty} dv_x \int_{-\infty}^{+\infty} dv_y \int_{-\infty}^{+\infty} dv_z\, \frac{1}{2} m_e v^2 f_e.$$

Questi integrali sono chiamati «momenti» (rispettivamente di ordine 0, 1 e 2) della funzione di distribuzione f_e. Ci si può chiedere allora se le grandezze n_e, $\mathbf{u}_e$, T_e (e quelle corrispondenti per gli ioni) che sono sostanzialmente quelle della magnetofluidodinamica, obbediscano ad equazioni simili, corrispondenti alla conservazione del numero delle particelle, della loro quantità di moto e della loro energia. Queste equazioni si possono facilmente ottenere mediante opportune integrazioni dell'equazione di Vlasov; ma si trova che esse non costituiscono un sistema chiuso, in quanto la derivata temporale dei momenti di ordine n dipende dai momenti di ordine $n+1$. Si può dire che l'equazione di Vlasov è equivalente a un sistema di infinite equazioni del tipo della magnetofluidodinamica. Solo in certe particolari approssimazioni è lecito trascurare i momenti del terzo ordine e chiudere un sistema per i momenti di ordine 0, 1 e 2 di ciascuna specie di particelle presenti; le equazioni che cosí si ottengono costituiscono una generalizzazione della magnetofluidodinamica ordinaria.

Rispetto alla magnetofluidodinamica, questo schema teorico ha un'altra differenza essenziale. Poiché è stato trascurato l'effetto delle collisioni, la direzione in cui fluisce il tempo è arbitraria; in altre parole, da una soluzione se ne ottiene automaticamente un'altra cambiando segno alla variabile temporale t e a tutte le componenti della velocità. Questa proprietà è la diretta conseguenza del fatto che le equazioni di moto per la particella singola godono anch'esse di questa simmetria rispetto all'inversione del tempo. Quando invece si fa l'ipotesi dell'equilibrio termico locale, la direzione in cui scorre il tempo è implicitamente fissata come quella nella quale il disordine (o l'entropia) aumenta. Quando s'inverte la freccia del tempo, si hanno processi in cui il caos molecolare diminuisce: una situazione teoricamente possibile, ma estremamente improbabile. Questa importante distinzione si vede anche dal fatto che nella descrizione di Vlasov non hanno posto i fenomeni di trasporto (per esempio, la conduzione termica), che ovviamente avvengono in senso opposto e non fisico quando si inverte il tempo.

La descrizione della dinamica del plasma in assenza di collisioni è assai piú complessa di quella con un modello fluido. Lo schema matematico a cui si è accennato ha dato luogo a straordinari sviluppi a partire dagli anni '60, con i quali è stata messa in evidenza una grandissima varietà di effetti inaspettati, di proprietà raffinate, di configurazioni geometriche peculiari. Si può dire, in maniera qualitativa, che un plasma senza collisioni gode di molti piú gradi di libertà di quando è descritto nell'approssimazione di un fluido ordinario; in altre parole il plasma è in grado di muoversi e deformarsi in un numero assai maggiore di maniere. In conseguenza di questo fatto è assai piú facile incontrare configurazioni instabili, in cui una perturbazione iniziale arbitrariamente piccola viene rapidamente amplificata, sino a stravolgere completamente l'equilibrio di partenza. Un'instabilità tipica di questo tipo ha luogo quando un fascio di elettroni attraversa il plasma, una situazione che non si può descrivere con un unico fluido. In questo caso, se la temperatura non è troppo alta, insorgono rapidamente oscillazioni elettrostatiche a frequenze dell'ordine della frequenza di plasma, che dànno luogo a uno stato turbolento. Le instabilità sono d'importanza cruciale in tutti i progetti di confinamento di un plasma: essi sono severamente condizionati da oscillazioni instabili con periodi piú piccoli del tempo di collisione, che spesso disperdono rapidamente il gas verso le pareti. A ciò si riesce ad ovviare solamente attraverso un'analisi precisa delle instabilità rilevanti e ricorrendo a condizioni di equilibrio assai complesse e con campi magnetici molto intensi. Come si vedrà nell'ultimo paragrafo, a questo problema è in parte dovuta la grande difficoltà incontrata nei programmi per realizzare in laboratorio la fusione nucleare controllata.

L'insorgenza di modi di oscillazione instabili in un plasma in assenza di collisione ha dato luogo a sviluppi teorici di grandissimo interesse, con i quali è stato possibile analizzare le interazioni che hanno luogo tra questi modi quando la loro ampiezza è sufficientemente grande. Due modi con frequenze ω_1 e ω_2 possono eccitare solamente i modi di frequenze $\omega_1 \pm \omega_2$; ciò corrisponde al principio di conservazione dell'energia in meccanica quantistica, dal momento che un fotone di frequenza ω ha un'energia $h\omega/2\pi$ (h è la costante di Planck); similmente, se le loro lunghezze d'onda sono λ_1 e λ_2 e le loro direzioni di propagazione $\mathbf{n}_1$ ed $\mathbf{n}_2$, i modi eccitati soddisfano sempre alla relazione

$$\frac{1}{\lambda}\mathbf{n}=\frac{\mathbf{n}_1}{\lambda_1}+\frac{\mathbf{n}_2}{\lambda_2},$$

corrispondente alla conservazione della quantità di moto in meccanica quantistica. Si può comprendere che l'insieme dei modi eccitati cosí selettivamente e i loro ritmi di crescita o di smorzamento, che si possono calcolare con precisione, diano luogo a una descrizione dettagliata degli stati turbolenti di un plasma, con svariate possibilità di applicazione.

Occorre sottolineare ancora che la differenza tra questi due schemi teorici

non sta tanto nel tipo di plasma considerato, ma piuttosto nella classe dei fenomeni che si vogliono descrivere. L'equazione di Vlasov, per esempio, non permette lo studio dei fenomeni di trasporto, che avvengono su scale di tempo assai piú lunghe dei periodi caratteristici delle oscillazioni menzionate. Si può rimediare a questa carenza senza ricadere nello schema fluido, troppo grossolano per i fenomeni piú interessanti, solo introducendo un terzo livello di descrizione matematica di cui si parlerà ora brevemente. L'equazione di Vlasov, come si è detto, descrive il moto delle particelle cariche sotto l'azione dei campi elettrici e magnetici esterni e di quelli generati collettivamente dalle particelle stesse; sono escluse le forze caotiche dovute alle singole collisioni. Questi due tipi di forze hanno dal punto di vista concettuale proprietà assai diverse: le prime sono deterministiche, cioè sono perfettamente fissate quando siano assegnate le forze esterne e la distribuzione di carica e di corrente elettrica nel plasma; la traiettoria di una particella è allora perfettamente determinata. Al contrario, poiché non è possibile predire l'istante e le modalità (direzione e intensità) di ciascuna collisione, queste generano necessariamente un'incertezza nella conoscenza del moto di ciascuna carica. Attorno alla traiettoria media si genera cosí, per modo di dire, una «nube di probabilità», che rappresenta la zona ove ci si aspetta di trovare la particella a ogni istante. Ragionando con lo spazio della velocità, ogni collisione produce una piccola, ma imprevedibile variazione della velocità della particella, come un ubriaco che si aggira senza meta attorno al punto di partenza; questo processo si chiama «cammino a caso» (*random walk*) e, se ha luogo nello spazio ordinario, produce una diffusione delle particelle da un zona piú popolata a una meno popolata. Nel caso presente si ha a che fare con una diffusione nello spazio delle velocità che tende a renderne la distribuzione isotropa e boltzmanniana.

Nella sua formulazione piú semplice questo effetto è descritto aggiungendo all'equazione di Vlasov un secondo membro che contiene le derivate prime e seconde di f rispetto alla velocità ed è anch'esso non-lineare rispetto alla f stessa. L'equazione che cosí si ottiene prende il nome, oltre che da Vlasov, da Fokker e Planck che per primi hanno dato una descrizione precisa dei fenomeni di diffusione. Essa, naturalmente, non è piú invariante quando la freccia del tempo è invertita e presenta difficoltà matematiche notevolmente superiori a quella dell'equazione di Vlasov.

In chiusura di questo paragrafo, un'osservazione di principio. Ci si potrebbe domandare: qual è la giustificazione teorica rigorosa di questi diversi livelli di descrizione? Giacché si sanno scrivere esattamente le equazioni di moto di N particelle cariche interagenti, in quale senso quando N è assai grande esse sono approssimabili da una di queste equazioni? Questo grande problema teorico costituisce l'oggetto di una disciplina particolare, la «teoria cinetica», ed è stato lungamente studiato.

A che livello e in quale approssimazione da equazioni dinamiche perfettamente reversibili (quelle per le singole particelle) appare il fenomeno dell'irreversibilità? Per rispondere a queste domande sono stati usati metodi statistici estremamente raffinati. Si può però dire almeno che il grado di irreversibilità

dipende dalla minor raffinatezza della descrizione usata; in altre parole, se l'oggetto delle misure e della teoria sono grandezze in qualche senso mediate, esse si evolvono in maniera irreversibile. Per esempio, se invece di considerare la velocità istantanea di una particella ci si limita a studiarne la media temporale su un tempo piú lungo della durata di una collisione tipica, ma piú corto delle scale temporali che interessano, tale media obbedisce a equazioni irreversibili. Quando un elettrone subisce una violenta collisione con un protone la sua energia cinetica $m_e v^2/2 \approx kT$ è dello stesso ordine di quella potenziale e^2/r, per cui la loro distanza è all'incirca $e^2/m_e v^2$; la durata di una collisione è allora $t_c = e^2/m_e v^3$. Oltre a questo si sono incontrati altri due tempi caratteristici: il periodo di plasma ω_p^{-1} e il tempo medio tra due collisioni $\tau \approx \omega_p^{-1} n d^3$. Un piccolo calcolo algebrico mostra che $\omega_p t_c \approx 1/nd^3$ (trascurando fattori numerici dell'ordine dell'unità che a questo scopo sono irrilevanti). Si vede quindi che, quando il numero nd^3 è assai grande, come di solito accade, $\tau \gg \omega_p^{-1} \gg t_c$. In queste due disuguaglianze, in genere verificate assai bene, sta il segreto della descrizione statistica del plasma. La prima assicura che le collisioni hanno scarsa importanza (per frequenze dell'ordine di ω_p); la seconda garantisce la validità del modello statistico, in cui si possono trascurare i dettagli delle singole collisioni e si privilegia una direzione del tempo.

3. *Plasma cosmico.*

Si può far passare una sostanza dallo stato liquido a quello aeriforme non solo portando la temperatura al di sopra dal punto di ebollizione, ma anche diminuendo la pressione. Ciò accade anche nel passaggio da aeriforme a plasma e in questo caso ha una semplice giustificazione fisica. La ionizzazione è prodotta da una collisione tra un atomo neutro e, ad esempio, un elettrone la cui energia cinetica relativa è maggiore dell'energia di ionizzazione; ma per il processo inverso – la ricombinazione – non è sufficiente che un elettrone passi vicino a un nucleo per venire catturato. Le forze elettrostatiche, difatti, non mutano l'energia totale del sistema; mentre la ricombinazione comporta necessariamente la diminuzione dell'energia da valori positivi (sistema slegato) a valori negativi (sistema legato). La ricombinazione può avvenire quando, per esempio, tre particelle interagiscono contemporaneamente, in maniera tale che la terza, aumentando la sua energia, diminuisca la quota spettante alle altre due. Ciò si può verificare, per esempio, quando un elettrone urta un protone in prossimità di un altro nucleo. Ciò mostra che, anche se l'evento è energeticamente possibile, la ionizzazione è molto piú probabile della ricombinazione; e questa differenza è tanto piú forte quanto piú rarefatto è il gas e, quindi, piú rare sono le collisioni a tre corpi. In questo caso l'equilibrio tra le due reazioni di ionizzazione e ricombinazione si sposta verso la prima. Per questa ragione i gas estremamente rarefatti che s'incontrano negli spazi interplanetari, interstellari e intergalattici sono spesso completamente ionizzati, anche se la loro temperatura corrisponde a un'energia cinetica media assai inferiore all'energia di ionizzazione.

Inizia qui una carrellata sui plasmi che s'incontrano in astrofisica dallo spazio interplanetario. Esso è riempito da un flusso di protoni, elettroni e nuclei d'elio emessi dal Sole; quando questo «vento solare» passa in prossimità della Terra contiene all'incirca 10 elettroni per cm^3, ha una velocità radiale di 400-500 km/sec e una temperatura di circa 100 000 gradi. In vicinanza del Sole esso è piú denso e piú caldo e costituisce la corona solare, visibile anche ad occhio nudo in occasione di un'eclissi sotto forma di grandi protuberanze luminose che partono dalla superficie del Sole. L'origine di questo grandioso fenomeno e della sua elevatissima energia termica non è ancora completamente chiarita, ma ha certo a che fare con lo stato assai turbolento dello strato esterno del Sole. Quando il vento solare investe il campo magnetico della Terra ne comprime il fronte e ne distende la parte posteriore in una lunga coda con un meccanismo che è sostanzialmente descritto dal complesso equilibrio tra la pressione magnetica interna e la pressione del fluido (fig. 5). Le particelle del vento solare possono rimanere intrappolate sulle linee di forza del campo magnetico terrestre che, essendo piú intenso in basso, si comporta come uno specchio magnetico (fig. 6). Le linee di forza che arrivano sulla Terra in prossimità dei poli sono aperte e direttamente

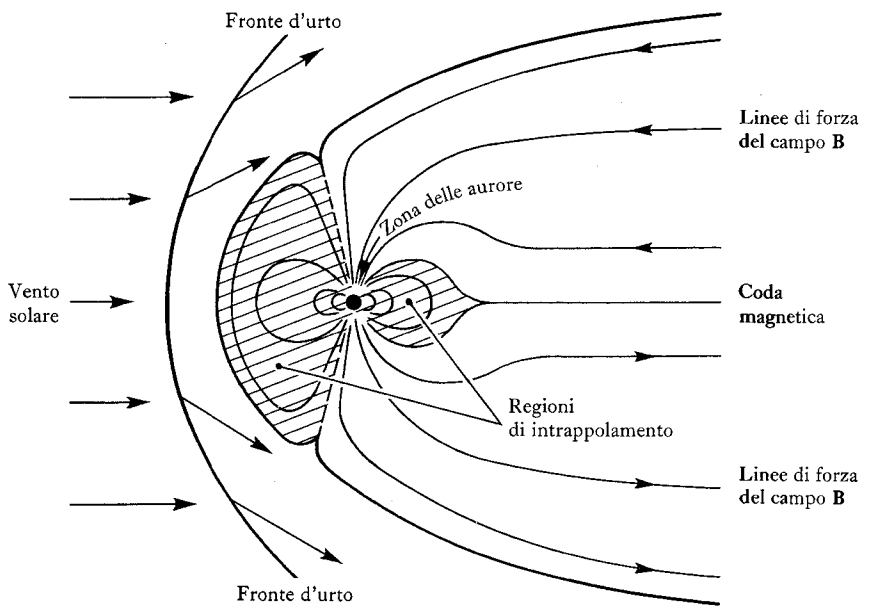


Figura 5.
Struttura della magnetosfera terrestre. Il vento solare viene deflesso dal campo magnetico della Terra in corrispondenza a una superficie chiamata «fronte d'urto». Le particelle del vento solare possono rimanere intrappolate nelle regioni tratteggiate; ai confini di queste regioni, in prossimità dei poli, esse possono precipitare nell'atmosfera, producendo le aurore boreali.

accessibili al vento solare; su di esse si può produrre un flusso di particelle che eccita le molecole dell'alta atmosfera producendo le aurore boreali. Al di sopra di 80 km l'atmosfera, estremamente rarefatta, è ionizzata e svolge un ruolo fondamentale nella propagazione delle onde radio (ionosfera).

Anche il campo magnetico della Terra è dovuto a un particolare tipo di plasma presente nel suo nucleo. Esso è costituito da silicati e metalli; questi ultimi, per la loro particolare struttura microscopica, sono in grado di generare elettroni liberi. Questa è la ragione per cui allo stato solido e liquido sono ottimi conduttori dell'elettricità. All'interno della Terra l'elevatissima pressione dovuta al peso degli strati superiori ne aumenta la temperatura al di sopra del punto di fusione; il liquido cosí generato si comporta come un plasma e obbedisce alle leggi della magnetofluidodinamica. Su di esso agiscono, oltre alla forza di gravità, le forze apparenti dovute alla rotazione terrestre (in particolare, quella centrifuga). Analisi assai complesse di questo sistema hanno mostrato che è possibile che venga spontaneamente generato, a spese dell'energia rotatoria, un campo magnetico parallelo all'asse, ma di verso arbitrario, proprio come indica l'evidenza geologica circa l'alternarsi tra polo Nord e polo Sud (effetto «dinamo»).

Le stelle ordinarie, come il Sole, sono formate prevalentemente da idrogeno ed elio completamente ionizzati e mantenuti a temperatura di decine di milioni di gradi dalle reazioni nucleari di fusione che avvengono nel loro interno. Quando nuclei leggeri – come l'idrogeno – si uniscono a formare nuclei piú pesanti – come l'elio –, viene liberata una grandissima quantità di energia; ma perché ciò avvenga è necessario che essi collidano con impulso sufficiente a superare la forza di repulsione elettrostatica che si esercita tra due cariche dello stesso segno.

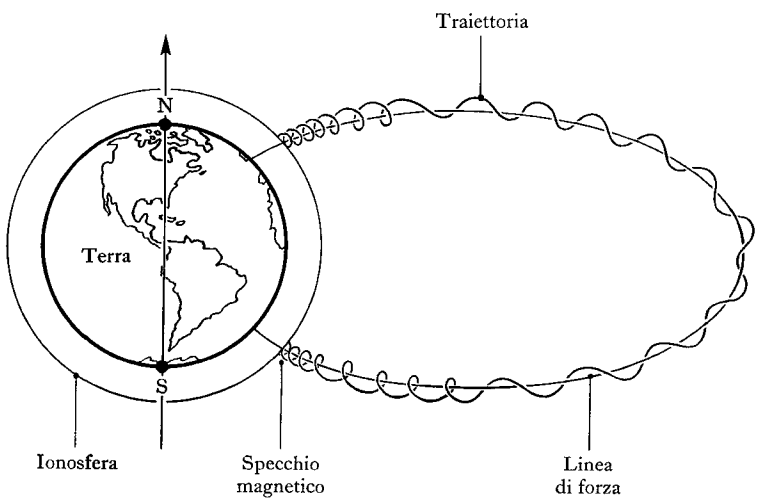


Figura 6.
L'intrappolamento delle particelle nel campo magnetico terrestre.

Se i due nuclei riescono ad avvicinarsi a distanze sufficientemente piccole (dell'ordine di 10^{-13} cm) le forze nucleari (attrattive) hanno il sopravvento e la loro fusione è possibile. Per questo occorre quindi che siano verificate due condizioni: 1) la temperatura del gas deve essere sufficientemente alta, e 2) la probabilità che un nucleo subisca una collisione di questo tipo prima di uscire dalla massa deve essere apprezzabile. Si vedrà nell'ultimo paragrafo che il raggiungimento di queste due condizioni costituisce la premessa fondamentale per realizzare in laboratorio la fusione controllata dell'idrogeno al fine di produrre energia; nel Sole ciò accade sostanzialmente per la sua grande massa, che richiede un'elevata temperatura per compensare la forza gravitazionale all'interno. L'energia cosí generata fluisce lentamente verso l'esterno e viene dispersa nello spazio sotto forma di energia radiante; il Sole è quindi una macchina che trasforma energia nucleare in energia elettromagnetica. A mano a mano che la fusione procede, si forma all'interno un nucleo di elio che aumenta lentamente il suo raggio a spese dell'idrogeno all'esterno; quando questo è esaurito si innescano reazioni piú complicate che, fondendo insieme nuclei di elio, generano specie nucleari piú massive.

Non è questo il luogo adatto per affrontare il problema centrale dell'astrofisica, cioè la struttura e l'evoluzione delle stelle; basti però notare che il plasma da cui sono costituite è certamente in equilibrio termodinamico, poiché il cammino che una particella percorre tra una collisione e l'altra è estremamente piú piccolo delle loro dimensioni.

Da ultimo, occorre parlare brevemente di un altro plasma cosmico di fondamentale importanza per l'evoluzione di tutto l'universo. Le stelle sono raggruppate in grandi sistemi tenuti anch'essi insieme dalla loro forza gravitazionale: le galassie. Esse contengono, oltre che decine o centinaia di miliardi di stelle, anche del gas – in parte ionizzato – in seno al quale esse si sono formate. La scoperta cruciale della cosmologia moderna sta nel fatto che le galassie sono animate da un rapito moto di recessione l'una dall'altra, la cui velocità aumenta all'aumentare della distanza. Si può concludere da ciò che in un tempo passato sufficientemente lontano le galassie non esistevano ancora e la loro materia era contenuta in un piccolo volume spaziale; a quell'epoca le stelle non erano nate e l'universo era formato da idrogeno neutro, con una percentuale di elio. Se s'immagina di seguire ancora all'indietro l'irresistibile moto di espansione cosmica, si troverà che questo gas, all'aumentare della sua densità e della sua temperatura, diventa ionizzato: a quel punto le dimensioni dell'universo erano circa un millesimo di quelle attuali e la sua età di circa un milione d'anni, da confrontare con l'età presente di quindici miliardi d'anni. In questa fase non si trova alcuna delle complesse e numerose forme di aggregazione che costituiscono il cosmo come noi lo vediamo: non c'è altro che un insieme perfettamente caotico di fotoni, di elettroni, protoni e nuclei d'elio animati da un vivacissimo moto di agitazione termica che impediva la formazione di qualsiasi aggregato stabile. Da questo caos primordiale, per raffreddamento ed espansione, sotto l'azione potente ed irresistibile della forza di gravità, si sono formate prima le galassie e poi, in seno ad esse, le stelle ed i pianeti quali ci appaiono nella loro infinita varietà e differenziazione. La maniera e i tempi in cui dal plasma primigenio si sono condensate le forme ordinate dell'universo costituisce un affascinante problema della cosmologia, in molti aspetti non ancora chiarito completamente.

4. *Energia dal plasma.*

Si è già detto che il motivo principale dello straordinario sviluppo sperimentale e teorico della fisica del plasma a partire dagli anni '60 sono stati i grandi investimenti di denaro e di energia umana che i paesi industrializzati hanno coscientemente dedicato a questo campo in vista della possibile realizzazione della fusione nucleare controllata. Al pari della fisica spaziale (nata dai missili balistici strategici) e dell'elettronica veloce (sviluppata durante l'ultima guerra per l'intercettazione degli aerei nemici col radar), essa mostra un altro importante esempio di processo di evoluzione scientifica invertito rispetto a quello comunemente conosciuto: qui la scoperta teorica non precede l'applicazione ma, piuttosto, gli aspetti di base e fondamentali vengono sviluppati sotto la spinta potentissima di una prospettiva concreta di un utile economico. Anche in questo caso l'origine lontana di questo progetto è militare. Il 1° marzo 1954 l'esplosione sperimentale prodotta dagli Stati Uniti a Bikini mostrò la possibilità di realizzare sulla Terra grandissime quantità di energia sfruttando la stessa reazione nucleare di fusione che ha luogo nel Sole con la bomba H. L'idea di realizzare in laboratorio in forma piú o meno controllata lo stesso processo e di utilizzare a fini civili l'energia cosí liberata si presentò subito e fu attivamente perseguita all'inizio in laboratori militari. Apparve però ben presto chiaro che il problema del controllo della reazione richiedeva conoscenze sperimentali e teoriche del plasma – in cui essa doveva necessariamente aver luogo – assai piú approfondite di quelle allora disponibili. Nel 1958 negli Stati Uniti a questo progetto fu tolto il vincolo di segretezza e da allora i fondi messi a disposizione in tutto il mondo a questo fine aumentarono costantemente, senza però raggiungere una soluzione definitiva. La crisi energetica ha reso ancora piú pressante questo progetto mondiale: difatti, se esso venisse realizzato, si utilizzerebbe come combustibile l'idrogeno dell'acqua.

Conviene precisare le due condizioni indicate nel paragrafo precedente per la realizzazione della fusione. La soglia di temperatura necessaria dipende un po' dalla reazione scelta, ma si può prendere come valore tipico cento milioni di gradi. Il tempo di confinamento t_{conf} è, naturalmente, funzione della densità, misurata dal numero degli elettroni per centimetro cubo, n; con precisione il prodotto nt_{conf} deve essere superiore a 2×10^{14} sec cm^{-3}. Per realizzare la seconda condizione si possono seguire due vie. Quella piú naturale e intrapresa per prima consiste nell'aumentare t_{conf} il piú possibile, mantenendo per n valori tipici dei plasmi ordinari prodotti in laboratorio ($10^{14} - 10^{15}$ cm^{-3}). All'inizio i sistemi di confinamento magnetico nei tori (fig. 3) e nelle macchine a specchi (fig. 4) incontrarono grossi insuccessi: il plasma, non appena prodotto entro i contenitori magnetici, scompariva subito. Iniziò allora la caccia e la cura delle diverse e sempre

piú raffinate forme d'instabilità che erano responsabili di questo fenomeno; allo stesso tempo si costruirono macchine sempre piú grosse per alleviare il problema aumentando le dimensioni del plasma. Questa via alla fusione è oggi attivissima e se ne vedrà piú avanti qualche esempio.

La seconda maniera consiste nell'abbandonare ogni tentativo di confinamento esterno e di lasciare che t_{conf} abbia il valore – brevissimo – che corrisponde all'espansione naturale di una piccola sfera di un composto d'idrogeno alla temperatura richiesta. Occorre però allora aumentarne la densità a valori superiori a quelli della materia solida. È necessario quindi disporre di un sistema per il riscaldamento e la compressione della sfera in tempi piú brevi di t_{conf}. Nella bomba H ciò è realizzato mediante l'esplosione attorno alla sfera di un guscio formato da materiale fissile che viene fatto esplodere come una bomba atomica ordinaria; in laboratorio occorre usare maniere piú delicate; per esempio, si può irraggiare la sfera con parecchi fasci di onde elettromagnetiche di altissima potenza generati da un sistema di laser. Questa radiazione viene assorbita dal materiale, che cosí viene rapidamente riscaldato e compresso. La figura 7 mostra una macchina di questo tipo – chiamata «Shiva» – in corso di costruzione nel laboratorio di Livermore, in California; essa costerà piú di 200 milioni di dollari. La riuscita di questo tentativo è legata alla disponibilità di laser estremamente potenti e alla realizzazione di dispositivi estremamente elaborati dal punto di vista tecnologico. Teoricamente, i complessi fenomeni che hanno luogo quando del materiale solido viene irradiato con altissimi flussi luminosi non sono ben conosciuti, soprattutto perché essi variano con grande rapidità. Sono assai utili a questo riguardo «esperimenti numerici» in cui le equazioni che governano il problema vengono risolte caso per caso su un grande elaboratore elettronico.

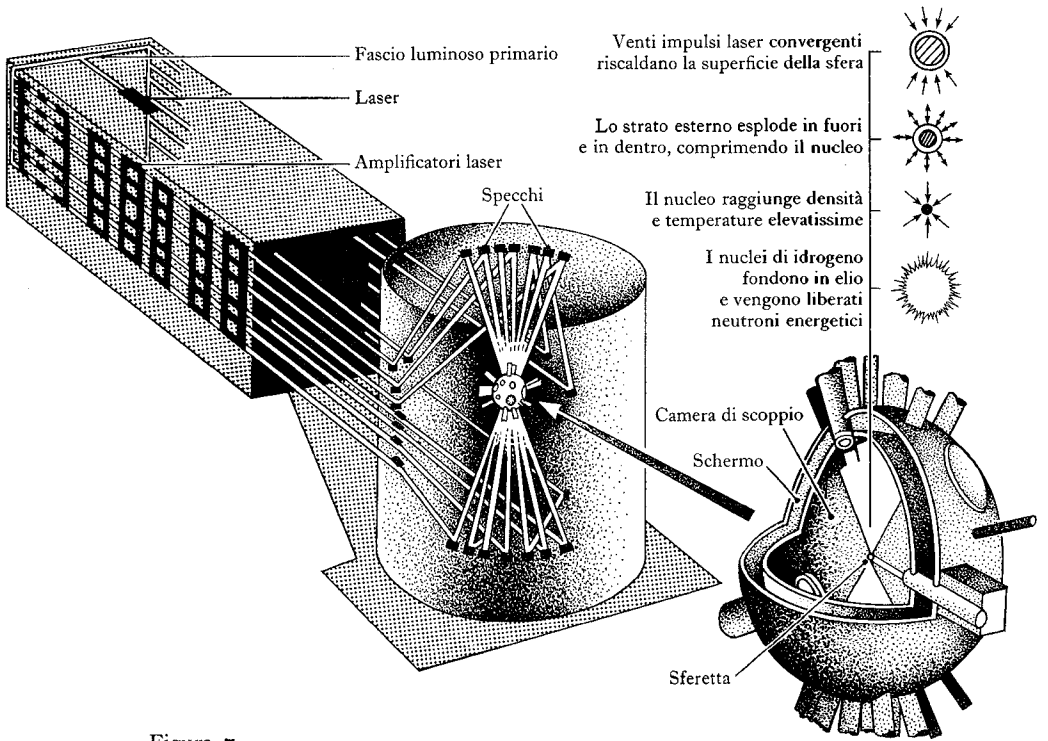


Figura 7.

«Shiva», la macchina in progetto nel laboratorio di Livermore per produrre fusione in una sferetta di un composto d'idrogeno mediante l'irradiazione simultanea di venti fasci laser.

Tornando alla prima via per realizzare la fusione nucleare controllata, si deve anzitutto parlare delle macchine toroidali, il cui principale tipo prende il nome russo di Tokamak, dal nome del dispositivo che fu costruito per la prima volta negli anni '60 nell'Istituto per l'energia atomica Kurchatov di Mosca. Il plasma nel Tokamak è soggetto a due campi esterni: un campo magnetico («toroidale») le cui linee di forza sono chiuse ad anello lungo il toro; e un campo elettrico diretto nella stessa direzione. Quest'ultimo è prodotto mediante l'effetto di induzione magnetica, cioè variando il flusso magnetico che il plasma abbraccia; per far questo attorno al plasma è avvolto un nucleo di ferro che porta un avvolgimento primario di spire. Allorché in queste ultime viene fatta passare una corrente variabile, il plasma si comporta come il circuito secondario di un trasformatore e diventa sede di un campo elettrico toroidale. La corrente cosí generata ha due scopi: anzitutto, essa scalda il plasma – che ha una resistenza residua – per effetto Joule; in secondo luogo, essa genera un campo magnetico secondario, le cui linee di forza sono cerchi concentrici all'asse del toro (campo «poloidale»). Questo campo esercita un importante effetto stabilizzante sul plasma (fig. 8).

Un altro importante schema di confinamento è, come si è detto, lo specchio magnetico; esso però non è di per se stesso sufficiente, perché una frazione delle particelle riesce sempre a sfuggire dalle estremità se la loro velocità forma con il campo magnetico un angolo troppo piccolo. Questo problema è stato oggetto di molte ricerche; una possibile soluzione è sfruttata in un grosso progetto del laboratorio di Livermore (cfr. fig. 9).

Mentre la realizzazione della condizione 2) per la fusione può, in linea di principio, venire raggiunta con un plasma sufficientemente esteso, il suo riscaldamento alle temperature necessarie per la fusione incontra degli ostacoli. Ovviamente questo scopo non può essere raggiunto mediante la conduzione termica ordinaria, che richiede una sorgente a temperatura superiore; occorre invece iniettare rapidamente grandi quantità di energia e fare in modo che essa venga distribuita tra i nuclei sotto forma di energia cinetica. Il riscaldamento separato dagli elettroni, per esempio, costituisce solo una perdita, poiché essi non partecipano al processo di fusione. Il riscaldamento dei nuclei comporta necessariamente un disordinamento dell'energia iniettata per effetto delle collisioni che essi subiscono; esso sarà tanto piú rapido quanto piú piccolo è il tempo τ che, in media, intercede tra due collisioni successive. Il problema sta nel fatto che τ non è una costante, ma dipende esso stesso dalla temperatura T; piú precisamen-

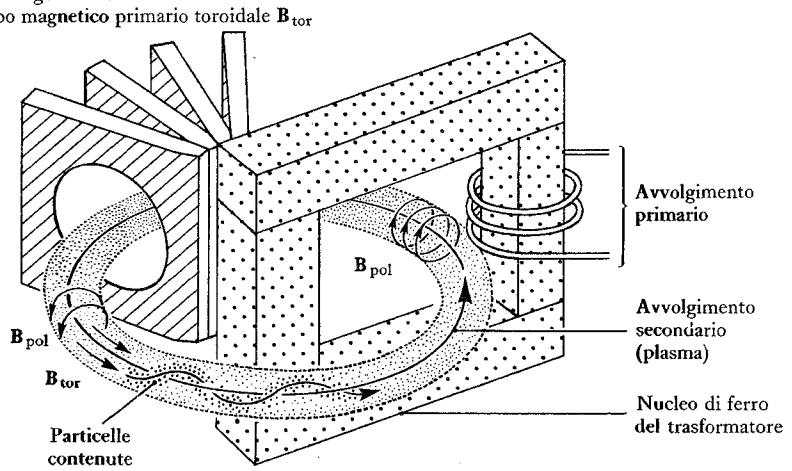


Figura 8.

Schema di principio di un Tokamak. Oltre al campo magnetico toroidale, vi è anche un campo secondario «poloidale» le cui linee di forza sono cerchi concentrici con l'asse del toro. Esso è generato dalla corrente toroidale che fluisce nel plasma, in quanto esso è il circuito secondario di un trasformatore. Le particelle vengono cosí confinate sulle linee di forza.

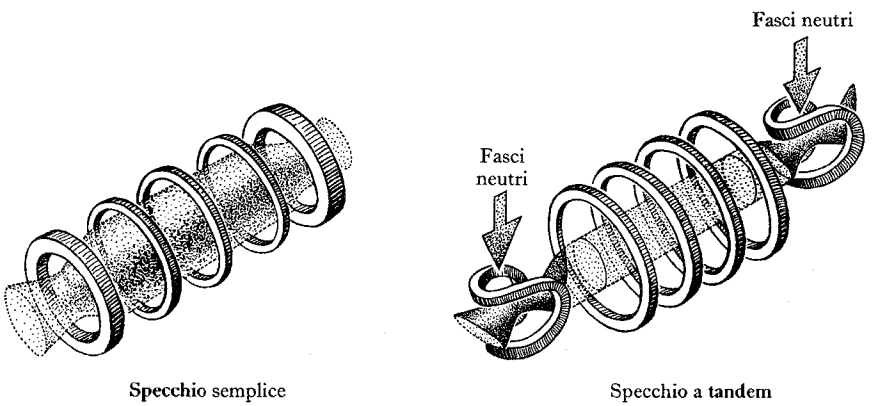


Figura 9.

Un sistema per chiudere le estremità di una bottiglia magnetica. Per mezzo degli speciali avvolgimenti a «palla da tennis» vengono create alle estremità delle configurazioni magnetiche speciali in cui il campo ha un minimo nel centro. In esse vengono iniettati dei fasci di gas neutro che, dopo essere stato ionizzato, rimane confinato sotto forma di plasma a densità relativamente alta; esso funziona come una specie di tappo agli specchi magnetici.

te, è proporzionale a $T^{3/2}$. Allora quando T aumenta, τ aumenta anch'esso, per cui il riscaldamento diventa sempre piú difficile e lento a mano a mano che il plasma viene scaldato. Tra i molti sistemi di riscaldamento sperimentati, oltre a quello basato sull'effetto Joule e sulla resistenza elettrica del plasma, va citata la possibilità di usare un intenso fascio di microonde e l'iniezione nella camera di un fascio di atomi neutri assai energetici che, traversando il campo magnetico, vengono ionizzati e subito dopo intrappolati.

La realizzazione di queste macchine, e ancor piú quelle che saranno necessarie nel futuro per estrarre effettivamente energia utile dalla fusione, hanno fatto della fisica del plasma una scienza costosa e complessa, il cui progresso dipende da grandi dispositivi strumentali, con formidabili problemi d'ingegneria e tecnologia avanzata. Questi progetti richiedono fondi elevati, grande coordinazione tra tecnici e scienziati di molti rami e un tempo notevole per la loro realizzazione; essi sono caratterizzati da una notevole rigidità, nel senso che le caratteristiche di queste macchine devono essere fissate molto tempo prima del loro uso effettivo ed è difficile mutarle secondo nuove prospettive o nuove concezioni.

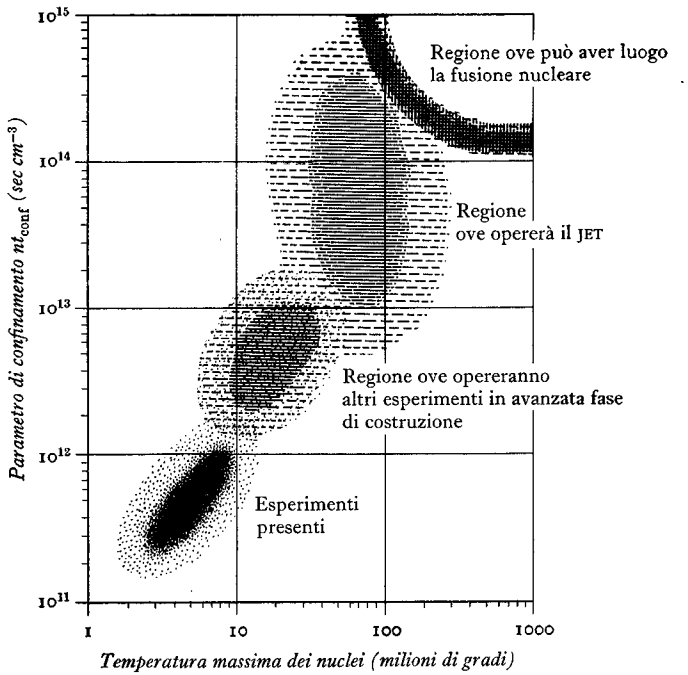


Figura 10.

Valori tipici dei due parametri cruciali per la fusione nucleare raggiungibili dalle macchine toroidali.

L'associazione europea dell'Euratom (di cui fa parte anche l'Italia) ha iniziato a costruire nel laboratorio di Culham, in Inghilterra una grande macchina toroidale (il JET). Altri tori di caratteristiche simili sono in costruzione o in progetto negli Stati Uniti, nell'Unione Sovietica e in Giappone. La costruzione del JET richiederà circa cinque anni e verrà a costare circa 150 miliardi di lire. Essa porterà abbastanza avanti nel soddisfacimento delle due condizioni necessarie per la fusione controllata, senza però raggiungerla (fig. 10). Per le macchine a specchi, la figura 11 mostra la concezione generale del progetto del laboratorio di Livermore. Naturalmente oltre a queste grandi realizzazioni ve ne sono altre minori. Tra di esse va ricordato il Tokamak costruito a Frascati dal Comitato nazionale per l'energia nucleare, che ha già dato interessanti risultati.

La realizzazione della fusione nucleare controllata in laboratorio è ancora lontana sia per le difficoltà tecnologiche sia per i problemi di fisica di base ancora da risolvere. Non si sa ancora qual è il sistema piú promettente: come si è visto,

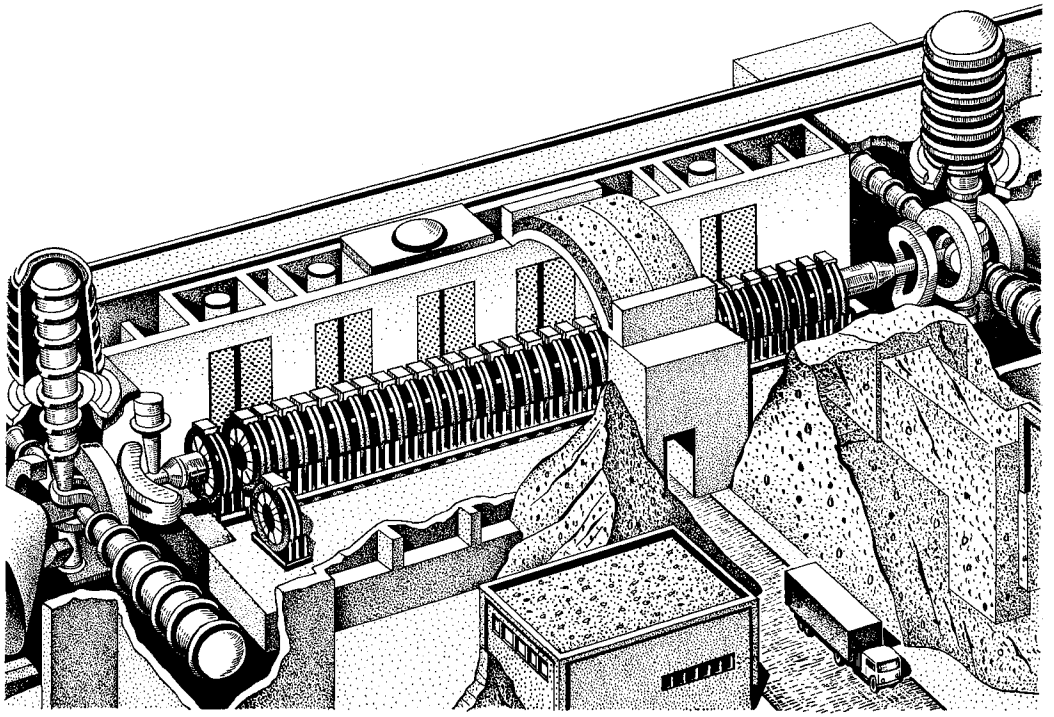


Figura 11.

Il quadro d'insieme della macchina in costruzione al laboratorio di Livermore (cfr. fig. 9). Si notino le dimensioni gigantesche in proporzione all'autocarro. All'estremità sinistra sono visibili gli iniettori di gas neutro e gli avvolgimenti a « palla da tennis ».

diversi tipi di macchine sono in competizione e, per esempio, solo nel 1995 il Dipartimento per l'energia degli Stati Uniti si ripromette di fare la sua scelta definitiva. Solo allora, nei piani americani, sarà possibile cominciare la costruzione di un prototipo che potrà dimostrare la fattibilità del progetto nel 2005. È già in fase di concezione una nuova macchina a confinamento toroidale – denominata INTOR (International Torus) –, che costituisce il successore del JET. Essa dovrebbe essere costruita negli anni '90 con una collaborazione tra Europa, Stati Uniti, Unione Sovietica e Giappone. [B. B.].

Lo studio della **fisica** del plasma, che costituisce il cosiddetto quarto stato della **materia**, ha assunto un'importanza sempre crescente sia per le sue implicazioni scientifiche sia per quelle economiche. Il **sole** e le stelle (cfr. **universo**) sono formati da plasma in equilibrio termodinamico (cfr. **equilibrio/squilibrio**) e quindi lo sviluppo della fisica dei plasmi ha notevoli implicazioni in **astronomia**, astrofisica, e nelle **cosmologie** (cfr. **ordine/disordine**); anche il nucleo della **terra** è costituito da plasma il cui movimento è all'origine del campo magnetico terrestre. Lo studio dei plasmi ha richiesto la costruzione di vari tipi di modelli (cfr. **modello**) nei quali vengono utilizzate le equazioni di conservazione (cfr. **conservazione/invarianza, forza/campo**), i concetti termodinamici (cfr. **energia, entropia, reversibilità/irreversibilità**) e quelli probabilistici (cfr. **caso/probabilità, distribuzione statistica**) della teoria cinetica dei gas. Le implicazioni economiche sono strettamente collegate alla possibilità di ottenere energia dal controllo della fusione termonucleare: si tratta di un obiettivo che richiede una collaborazione assai stretta fra **industria** e **scienza** (cfr. anche **tecnica, ricerca, teoria/pratica**) che viene incentivata dallo sperato utile economico (cfr. **profitto**). Si devono infatti ideare e costruire (cfr. **disegno/progetto, innovazione/scoperta, invenzione**) complessi e costosi meccanismi che controllino le instabilità (cfr. **stabilità/instabilità**) la cui **propagazione** può distruggere il plasma.

Propagazione

Un fenomeno fisico verificato in un punto è avvertito, piú tardi, in punti vicini, piú tardi ancora in punti piú lontani. Esistono due grandi tipi di fenomeni fisici: l'emissione di particelle da una parte, e la deformazione dei mezzi continui dall'altra. Si applicherà il termine 'propagazione' solamente a questi ultimi. Se ne osserva un tipico esempio lasciando cadere un sasso in uno stagno. Si nota il formarsi, sulla superficie dell'acqua, di cerchi concentrici col punto di caduta, che si allargano propagando la perturbazione originata dal sasso. Questa propagazione non è dovuta a trasporto di materia: infatti, un oggetto leggero che galleggi sull'acqua oscilla pressoché verticalmente, mentre le onde della superficie lo raggiungono e lo sorpassano. Un mezzo piú efficace per verificare che le onde non sono dovute a un arrivo periodico di particelle consiste nell'effettuare un'esperienza di interferenza: se si pone nello stagno una fila di pali, ciascuno di essi, quando viene interessato dalla perturbazione originale, si comporta come sorgente di una nuova perturbazione, e queste differenti perturbazioni interferiscono tra di loro rafforzandosi, o annullandosi se sono in senso opposto: se si trattasse di un insieme di particelle, alcune sarebbero riflesse dai pali, altre continuerebbero il loro cammino e non si avrebbe interferenza. Si dice che si ha «propagazione per onde». In fisica matematica, la superficie dello stagno che varia con il tempo viene definita onda, e si definisce fronte d'onda il cerchio di discontinuità che, in un dato istante, separa la regione non perturbata dalla regione interessata dalla perturbazione.

I fenomeni oscillatori (come le increspature sulla superficie dell'acqua) che si propagano nel tempo sono molto numerosi: il suono, la luce (nonostante questa non abbia supporto materiale) sono da annoverare in questa categoria. In via del tutto generale, si definisce onda qualsiasi fenomeno (non necessariamente periodico) che si propaghi senza essere direttamente legato a trasporto di materia; questo significa, a un approccio superficiale, che vi si può localizzare un segnale che permetta di seguirne lo spostamento nel tempo, malgrado le possibili deformazioni. Si partirà dalla legge di propagazione piú semplice, quella delle piccole oscillazioni di una corda tesa, per arrivare alla piú complessa, quella dei misteriosi solitoni.

1. *Propagazione di un'oscillazione lungo una corda tesa.*

Si consideri una corda elastica tesa AB di lunghezza ℓ (fig. 1) e s'imprima bruscamente con la mano, all'estremità A, un piccolo spostamento da A in A', riportandola poi da A' in A. L'elongazione y_A del punto A varia in funzione del tempo t seguendo la legge rappresentata nella figura 2.

L'esperienza mostra che lo spostamento impresso al punto A si propaga man mano con velocità costante v, che dipende solo dalla natura della corda e dalla

sua tensione, e non dipende dalla legge di spostamento (supposto piccolo) del punto A; il moto di un punto M situato a una distanza x da A sulla curva riproduce quello del punto A dopo un tempo $\tau = x/v$ (cfr. fig. 3).

Se si traccia un grafico dello stato della corda all'istante t con $(\ell/v) > t > t_2$, poi all'istante $t' = t + \vartheta$, si ha che la curva rappresentante le deformazioni si è spostata di $v\vartheta$.

Se $y = f(t)$ è l'equazione del moto del punto iniziale A, l'elongazione del punto M all'istante t è quella che aveva il punto A nell'istante $t - \tau$, $\tau = x/v$. L'equazione del moto di M è dunque

$$(1) \qquad y = f(t-\tau) = f\left(t - \frac{x}{v}\right).$$

Questa è l'equazione generale di un'oscillazione che si propaga lungo l'asse ox, con velocità v. L'elongazione (1) è una funzione delle due variabili x e t, e si definisce onda progressiva. Rappresenta il moto della corda fino a quando l'oscillazione non interessa l'estremità B, cioè per $t \leq (\ell/v)$.

Quando l'oscillazione interessa l'estremità B, supposta fissa, si nota che questa si riflette, cioè riparte in direzione contraria, con un'elongazione simmetrica e uguale velocità: è quindi rappresentata dall'equazione

$$y = -f\left(t - \frac{\ell}{v} + \frac{x-\ell}{v}\right) \qquad \text{se } \frac{2\ell}{v} > t > \frac{\ell}{v},$$

che è del tipo

$$(2) \qquad y = g\left(t + \frac{x}{v}\right)$$

dove g è una funzione di una variabile del tipo $g(u) = -f[u - (2\ell/v)]$. L'elongazione (2), è definita onda regressiva.

Entrambe le funzioni $f[t - (x/v)]$ e $g[t + (x/v)]$, supposte derivabili due volte, soddisfano l'equazione alle derivate parziali detta equazione di d'Alembert:

$$(3) \qquad \frac{\partial^2 y}{\partial t^2} - v^2 \frac{\partial^2 y}{\partial x^2} = 0$$

(si ha per esempio: $(\partial f/\partial t) = f'[t - (x/v)]$, $(\partial f/\partial x) = (-1/v) f'[t - (x/v)]$, $(\partial^2 f/\partial t^2) = f''[t - (x/v)]$, $(\partial^2 f/\partial x^2) = (1/v^2) f''[t - (x/v)]$).

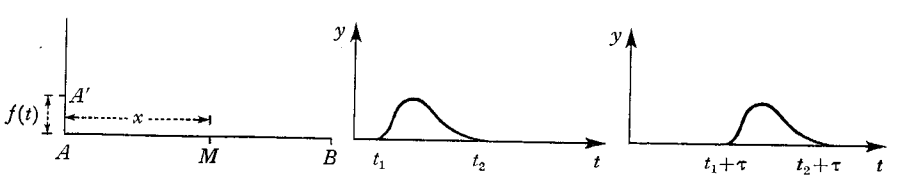


Figure 1-3.

Durante le oscillazioni di una corda tesa (fig. 1) lo spostamento del punto A (fig. 2) viene riprodotto dal punto M (fig. 3) dopo un tempo $\tau = x/v$.

Si può ottenere anche a priori questa equazione, scrivendo che l'accelerazione di ogni punto della corda è proporzionale alla risultante delle forze elastiche che agiscono su di esso in seguito alla piccola deformazione.

È facile dimostrare che la soluzione generale dell'equazione (3) è

$$(4) \qquad y = f\left(t - \frac{x}{v}\right) + g\left(t + \frac{x}{v}\right)$$

dove f e g sono funzioni arbitrarie, o anche, se si rimuove la condizione di essere due volte derivabili, distribuzioni (funzioni generalizzate).

Per questa dimostrazione, è sufficiente trasformare l'equazione (3) nell'equazione $\partial^2 y / \partial\xi\, \partial\eta = 0$ con il cambiamento di variabili: $\xi = t - (x/v)$, $\eta = t + (x/v)$.

S'interpreta l'equazione (4) dicendo che, per i casi piú generali, i piccoli moti della corda tesa risultano dalla sovrapposizione di due moti che si propagano in senso opposto con uguale velocità v.

Oggigiorno, l'equivalenza tra i risultati dedotti da questa soluzione generale e le soluzioni di problemi specifici costruite con sviluppi in serie di funzioni ottenute per «separazione di variabili», è stabilita in quadri funzionali ben precisi, ma si sono dovuti attendere gli analisti del XIX secolo affinché le dimostrazioni di convergenza delle serie risolvessero la controversia tra d'Alembert, che voleva considerare come soluzioni solo funzioni f e g analitiche, ed Eulero, che nei suoi sviluppi in serie ammetteva funzioni piú generali, pur dubitando della loro generalità.

2. *Programmazione di onde piane nei fluidi comprimibili e nei solidi elastici.*

L'equazione di d'Alembert, detta anche equazione delle corde vibranti, regola ancora la propagazione delle piccole oscillazioni nei fluidi perfetti comprimibili. Questo fatto fu stabilito da d'Alembert stesso nel 1752 nella memoria *Essai sur une nouvelle théorie de la résistance des fluides* da lui sottoposta al concorso indetto dall'Accademia delle Scienze di Berlino nel 1750; il premio, però,

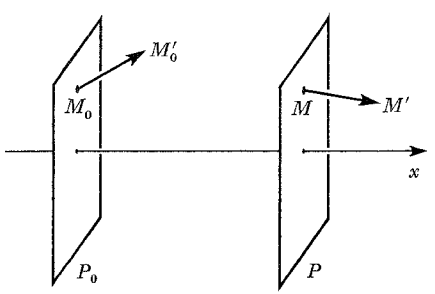


Figura 4.
Movimento per onde piane.

non fu attribuito a d'Alembert, in quanto la giuria gli rimproverò di non aver aggiunto alla sua memoria dei calcoli che confermassero l'accordo dei suoi risultati con l'esperienza. Infatti l'esposizione di d'Alembert del «moto per parti» era molto complessa. Verrà qui data, con linguaggio moderno, un'esposizione elementare che si applica anche al caso di solidi elastici, illimitati, omogenei e isotropi, in cui un piano P_0 è soggetto a un piccolo moto che lo lascia parallelo a se stesso.

Si supponga d'imporre a tutti i punti M_0 di un dato piano P_0 dei movimenti tali che, in qualsiasi istante t, i loro spostamenti M_0M_0' siano vettori uguali tra di loro. Se si definisce la posizione di un qualsiasi punto M del mezzo, con le sue coordinate x, y, z, rispetto a un triedro il cui asse ox è perpendicolare al piano P_0, il vettore spostamento MM' dipende solo dalle variabili x e t (fig. 4).

Si dice allora che il moto è un'onda piana, o che si propaga per onde piane. I differenti piani $x =$ costante, paralleli al piano P_0, in cui lo spostamento è lo stesso a un dato istante t, sono detti piani d'onda, la direzione ox della normale a questi piani è la direzione di propagazione.

Se il moto della «sorgente» (il piano P_0) inizia a un dato istante, per esempio $t = 0$, si definisce fronte d'onda all'istante t il piano piú distante in cui è avvertito questo moto.

2.1. Moti longitudinali e moti trasversali.

Si dice che un'onda piana è trasversale quando lo spostamento MM' di un punto qualsiasi è perpendicolare alla direzione di propagazione, quindi parallelo al piano d'onda P. Si dice che l'onda è longitudinale quando lo spostamento MM' è parallelo alla direzione di propagazione.

2.2. Propagazione di onde piane nei fluidi comprimibili.

In un fluido perfetto, cioè con viscosità trascurabile, non esiste nessuna forza che si oppone allo scorrimento dei diversi strati gli uni rispetto agli altri: non esiste propagazione di moti trasversali.

S'immagini, invece, in un fluido omogeneo e isotropo illimitato, di spostare leggermente nella direzione ox gli elementi di un piano P_0 perpendicolare a questa direzione: si genera allora una compressione (o una dilatazione) dello spazio compreso fra P_0 e un piano parallelo vicino P, che induce via via un moto del fluido. L'equazione cui soddisfa la funzione u che caratterizza questo moto può essere ottenuta nel modo seguente.

Si consideri una parte di liquido in quiete, compresa tra i piani P e P' di ascissa x e $x + dx$. Si supponga che in un istante t del moto studiato i piani in questione P e P' abbiano come ascissa $x + X$ e $x + dx + X + dX$: la corrispondente dilatazione di un elemento di volume del fluido (da α_0 a α) è

$$\vartheta = \frac{\alpha - \alpha_0}{\alpha_0} = \frac{dx + dX - dx}{dx}.$$

Passando al limite, $\vartheta = \partial X/\partial x$, dove la derivata parziale $\partial/\partial x$ ricorda che il calcolo è fatto in un istante t fissato. La derivata parziale $\partial X/\partial t$ è la velocità all'istante t della parte di fluido che, in quiete, aveva come ascissa x.

Tenendo presente il fatto che, per piccoli moti, il coefficiente di comprimibilità di un fluido è una costante χ, e cioè che la variazione di pressione $p = \pi - \pi_0$ è proporzionale, mediante il fattore χ, alla dilatazione ϑ, si ottiene il risultato $p = (-1/\chi)(\partial X/\partial x)$. Da qui, uguagliando quantità di accelerazione (parallela a ox) e forza applicata (risultante delle pressioni; cfr. fig. 5), risulta l'equazione

$$\rho \frac{\partial^2 X}{\partial t^2} = \frac{1}{\chi} \frac{\partial^2 X}{\partial x^2}$$

dove ρ è la densità del liquido. Si ritrova, quindi, l'equazione di d'Alembert

$$\frac{\partial^2 X}{\partial t^2} - v^2 \frac{\partial^2 X}{\partial x^2} = 0 \qquad \text{con } v = \frac{1}{\sqrt{\chi\rho}}.$$

Questa equazione è soddisfatta anche dalla velocità $u = \partial X/\partial t$, e dalla pressione p

$$\frac{\partial^2 p}{\partial t^2} - v^2 \frac{\partial^2 p}{\partial x^2} = 0 \qquad \frac{\partial^2 u}{\partial t^2} - v^2 \frac{\partial^2 u}{\partial x^2} = 0$$

(si ha $(\partial p/\partial t) = (-1/\chi)(\partial u/\partial x)$).

La soluzione generale $p = f[t-(x/v)] + g[t+(x/v)]$ è ancora la somma di un'onda progressiva e di un'onda regressiva che si propagano in senso contrario con uguale velocità v.

Un modello fisico dello schema precedente è dato da piccoli movimenti di un fluido perfetto, contenuto in un tubo cilindrico sufficientemente grande da poter trascurare l'influenza delle pareti. Si genera un'onda piana progressiva nel tubo, posizionando a una delle sue estremità un pistone P_0, al quale s'imprime un moto di piccola ampiezza rappresentato dalla legge $X = f(t)$: l'oscillazione si propaga con velocità v e il moto di un punto M di ascissa x è dato da $u = f[t-(x/v)]$, dal momento che l'oscillazione non interessa l'altra estremità del tubo.

In pratica la viscosità e gli scambi di calore con le pareti non sono affatto

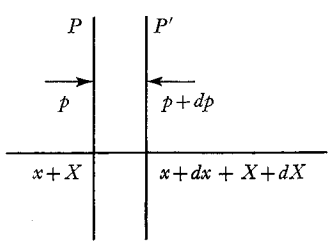


Figura 5.
Risultante delle pressioni sulla parte di liquido compresa fra i piani P e P'.

nulli, di modo che il segnale si smorza e la velocità di propagazione diminuisce nel tempo. Inoltre, la legge qui adottata di proporzionalità fra le dilatazioni e le variazioni di pressione è un'approssimazione valida solo quando queste variazioni rimangono piccole.

2.3. Misure della velocità del suono.

La propagazione del suono nell'aria obbedisce alla legge appena studiata: la velocità di propagazione è indipendente dal timbro del suono (fortunatamente per i melomani).

Newton ha dato per primo, per questa velocità, la formula $v = 1/\sqrt{\chi\rho}$. Egli suppose che la propagazione fosse isoterma (cioè a temperatura T costante) e utilizzò la legge dei gas perfetti, $\pi = R\rho T$, per determinare il coefficiente di comprimibilità χ; le misure sperimentali forniscono un risultato piú alto di quello ottenuto con il calcolo di Newton. Infatti le compressioni e le espansioni molto rapide che corrispondono alla propagazione del suono non sono isoterme ma isoentropiche, e il coefficiente χ deve essere calcolato, come fece osservare Laplace, partendo dalla legge isoentropica $\pi\rho^{-\gamma} =$ costante, dove γ è una costante, rapporto fra i calori specifici a pressione costante e a volume costante; in questo modo la velocità del suono è $v = \sqrt{\gamma\pi/\rho} = \sqrt{\gamma RT}$. La velocità del suono nell'aria dipende quindi solo dalla temperatura, e precisamente è proporzionale alla radice quadrata della temperatura assoluta. Questa velocità, nell'aria secca a zero gradi, è 330,75 m/s con l'approssimazione di circa mezzo decimetro.

2.4. Propagazione per onde piane nei solidi elastici.

Si consideri un solido elastico illimitato, omogeneo e isotropo, e s'imprima a un piano P_0 di questo solido, che in quiete ha ascissa x, una piccola traslazione di vettore $\boldsymbol{\delta}$ che lo porta su un piano parallelo P_0'. Il vettore $\boldsymbol{\delta}$ è la somma di due vettori: $\boldsymbol{\delta}_1$, parallelo all'asse ox ortogonale al piano P_0, e $\boldsymbol{\delta}_2$, ortogonale a questo asse: il moto risultante (se di ampiezza piccola) sarà la somma dei moti risultanti dallo spostamento $\boldsymbol{\delta}_1$ (detto longitudinale) e dallo spostamento $\boldsymbol{\delta}_2$ (detto trasversale). I movimenti longitudinali in una sbarra sono stati studiati nel 1817 e 1819 da Laplace e Poisson. Lamé, nel 1852, ha fornito un'elaborazione razionale del caso generale (a tre dimensioni), nelle sue *Leçons sur la théorie mathématique de l'élasticité*, che si basa in parte su lavori di Cauchy apparsi tra il 1815 e il 1830.

L'equazione soddisfatta dall'onda piana longitudinale si ottiene in modo analogo a quello prima utilizzato per i fluidi. Anche in questo caso la dilatazione è $\vartheta = \partial u/\partial x$, ma le forze applicate a una piccola parte del mezzo non si ottengono piú dalla formula relativa alla variazione di pressione e dilatazione, ma dalla legge dell'elasticità lineare, relativa a tensioni e deformazioni, $T = \lambda\vartheta + 2\mu a_1 = (\lambda + 2\mu)(\partial u/\partial x)$, dove λ e μ sono le costanti di Lamé relative al materiale. L'equazione soddisfatta da u è dunque ora

$$\frac{\partial^2 u}{\partial t^2} - (\lambda + 2\mu) \frac{\partial^2 u}{\partial x^2} = 0.$$

La velocità di propagazione di un'onda piana longitudinale è $\sqrt{(\lambda+2\mu)/\rho}$.

Un'onda piana trasversale corrisponde a un piccolo spostamento del piano P_0, caratterizzato da un vettore $\boldsymbol{\delta}_2$ parallelo a questo piano; una piccola parte PP' subisce uno scorrimento lungo se stessa, senza dilatazione. Si scelga per asse oz un asse parallelo a $\boldsymbol{\delta}_2$ e si indichi con w lo spostamento nella direzione oz, subito all'istante t da un punto M di ascissa x del solido. Anche la tensione esercitata sul piano P corrispondente è parallela ad oz, e ha come valore $Z = -(\partial w/\partial x)\mu$. Se ne deduce che l'equazione soddisfatta da w è

$$\rho\frac{\partial^2 w}{\partial t^2}-\mu\frac{\partial^2 w}{\partial x^2}=0.$$

La velocità di propagazione di un'onda piana trasversale è dunque $\sqrt{\mu/\rho}$.

L'onda piana corrispondente a una piccola oscillazione arbitraria $\boldsymbol{\delta}=\boldsymbol{\delta}_1+\boldsymbol{\delta}_2$ del piano P_0 è la somma delle onde piane generate dalle oscillazioni $\boldsymbol{\delta}_1$ e $\boldsymbol{\delta}_2$.

In natura non esiste un solido omogeneo indefinito, quindi non si possono verificare sperimentalmente in modo rigoroso le formule precedenti. Tuttavia queste dànno una buona interpretazione dei fenomeni osservati al momento della registrazione di sismi: si constata che i sismografi registrano sempre due serie di oscillazioni, i cui inizi corrispondono rispettivamente all'arrivo di onde longitudinali e onde trasversali. Per i sismografi non troppo lontani dall'epicentro del sisma (meno di 600-700 km), si trovano velocità di propagazione vicine a quelle teoriche, calcolate partendo dal coefficiente di elasticità delle rocce che si trovano nel suolo (in media $V_1 \simeq 5500$ m/s per le onde longitudinali, $V_2 \simeq 3200$ m/s per le onde trasversali).

Quando i sismografi sono piú lontani dall'epicentro del sisma, si registrano sempre due serie di vibrazioni, corrispondenti all'arrivo di onde longitudinali e di onde trasversali, ma le velocità di propagazione sono molto piú grandi. Una tale onda sismica ha percorso una traiettoria come P_0AP (cfr. fig. 6) che è penetrata molto profondamente nella Terra; se ne deduce (dato che la densità ρ cresce certamente con la profondità) che i coefficienti di Lamé, λ e μ, aumentano con la profondità. Si sono potute ottenere indicazioni molto precise sulla costituzione del globo terrestre (e piú recentemente della Luna) mediante lo studio

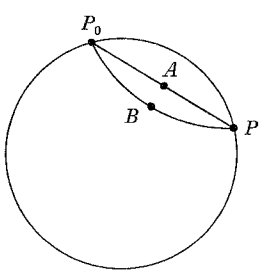


Figura 6.
Propagazione di onde sismiche nel mantello terrestre.

della legge di variazione della velocità di propagazione con la profondità. D'altra parte, questa variazione provoca una curvatura nella direzione di propagazione: l'oscillazione segue una traiettoria del tipo P_0BP (fig. 6). Per studiare la legge (teorica) di propagazione delle onde sismiche, come d'altra parte per le onde sonore e luminose, bisogna mettersi nella situazione realistica delle tre dimensioni spaziali.

3. *Equazione delle onde.*

L'equazione delle onde è la generalizzazione tridimensionale dell'equazione delle corde vibranti. Si scrive

$$\frac{\partial^2\varphi}{\partial t^2}-c^2\Delta\varphi=0 \tag{5}$$

dove c indica una costante e Δ il laplaciano, cioè in coordinate cartesiane

$$\Delta\equiv\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}. \tag{6}$$

L'equazione delle onde è fondamentale sia nello studio delle piccole perturbazioni dei fluidi comprimibili (studio definito acustico) e dei solidi elastici, sia nell'elettromagnetismo. È stata scritta fin dalla metà del XVIII secolo in relazione a problemi di aerodinamica, ma è stata studiata in modo completo solo molto piú tardi. Attualmente è considerata un prototipo di quelle equazioni iperboliche alle derivate parziali che governano la fisica.

Le equazioni generali della dinamica dei gas furono scritte da Eulero nel 1755, cioè poco dopo le memorie di d'Alembert sul moto per parti dei fluidi. L'esposizione di Eulero, definita da Lagrange «chiara e semplice», è sempre valida. Egli dimostra che, quando si trascurano la viscosità e gli scambi termici, le equazioni della dinamica dei gas sono, in assenza di forze esterne,

$$\frac{\partial\rho}{\partial t}+\operatorname{div}(\rho\mathbf{u})=0 \qquad \text{(conservazione della massa)}$$

$$\left(\frac{\partial\mathbf{u}}{\partial t}\right)+\mathbf{u}(\operatorname{grad}\mathbf{u})+\operatorname{grad}\pi=0 \qquad \text{(equazione vettoriale, conservazione del momento)}$$

dove ρ indica la densità, π la pressione e $\mathbf{u}$ il vettore velocità in un punto M all'istante t; se x_1, x_2, x_3 sono le coordinate cartesiane, e u_1, u_2, u_3 sono le componenti del vettore $\mathbf{u}$, si ha $\operatorname{div}(\rho\mathbf{u})=\sum_{i=1}^{3}(\partial/\partial x_i)(\rho\mathbf{u}_i)$ e $\mathbf{u}$ (grad $\mathbf{u}$) è il vettore di componenti $\sum_{j=1}^{3}u_j(\partial u_i/\partial x_j)$. Queste equazioni sono da completare con una equazione di stato, dipendente dalla natura del gas e relativa alla densità e alla pressione, $\pi=f(\rho)$ (ci si limita per semplicità ai movimenti isoentropici).

Se si considerano piccole perturbazioni, a partire da una posizione di equili-

brio in cui la pressione π_0 e la densità ρ_0 sono costanti, si pone: $\rho=\rho_0+r$ e $\pi=\pi_0+p$ e si trascurano tutti i termini che contengono prodotti di piccole quantità di r, p, $\mathbf{u}$ o di loro derivate. Si ottengono cosí le equazioni dette lineari, nella forma

$$\frac{\partial r}{\partial t}+\operatorname{div}(\rho_0\mathbf{u})=0$$

$$\rho_0\frac{\partial \mathbf{u}}{\partial t}+\operatorname{grad}p=0.$$

D'altra parte si ha, con la stessa approssimazione, $p=f'(\rho_0)r$ (con $f'=df/d\rho$), per cui: $\partial p/\partial t+\rho_0 c^2\operatorname{div}\mathbf{u}=0$ dove $c^2=f'(\rho_0)$. Se ne deduce che p soddisfa l'equazione delle onde

$$\frac{\partial^2 p}{\partial t^2}-c^2\operatorname{div}(\operatorname{grad}p)=\frac{\partial^2 p}{\partial t^2}-c^2\Delta p=0.$$

3.1. Propagazione per onde sferiche.

Si dice che una soluzione φ dell'equazione delle onde è un'onda sferica quando, in ogni istante t, ha lo stesso valore in tutti i punti di una qualsiasi sfera che abbia come centro un punto dato O fisso; in altre parole la funzione φ dipende solo dal tempo t e dalla distanza r da O al punto M variabile. In effetti si osservano onde sferiche, in un mezzo omogeneo e isotropo tridimensionale, quando la sorgente delle piccole perturbazioni può essere considerata puntiforme, oppure è, essa stessa, a simmetria sferica: per esempio, se si pone nell'aria, in un punto O, un pallone sferico che si gonfi e si sgonfi, si genera un'onda sferica.

Se l'incognita φ dell'equazione delle onde dipende solo da r e da t, applicando le regole elementari della derivazione di una funzione composta, si trova che φ soddisfa l'equazione

$$\frac{\partial^2\varphi}{\partial t^2}=c^2\left(\frac{\partial^2\varphi}{\partial r^2}+\frac{2}{r}\frac{\partial\varphi}{\partial r}\right).$$

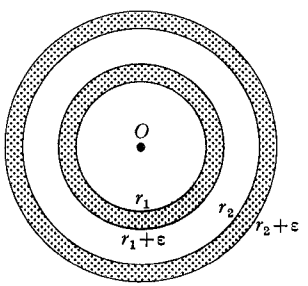


Figura 7.

Propagazione di un'onda sferica: il fronte d'onda si propaga con velocità c e la regione interessata dall'oscillazione ha lo spessore costante ε.

Considerando allora la funzione $\psi=r\varphi$, si vede che l'equazione precedente diventa

$$\frac{\partial^2\psi}{\partial t^2}-c^2\frac{\partial^2\psi}{\partial r^2}=0,$$

uguale all'equazione di d'Alembert (delle corde vibranti e delle onde piane), la cui soluzione generale è $\psi=f(ct-r)+g(ct+r)$. Quindi la soluzione generale dell'equazione delle onde sferiche è

$$\varphi(r,t)=\frac{f(ct-r)}{r}+\frac{g(ct+r)}{r}.$$

Il moto corrispondente a un'onda sferica può sempre essere considerato come la somma di due moti che si propagano in senso contrario: il primo, rappresentato dalla funzione $f(ct-r)/r$, si propaga nella direzione di r crescente, si allontana dall'origine O e viene detto onda divergente; il secondo, rappresentato dalla funzione $g(ct+r)/r$, si propaga verso l'origine O e viene detto onda convergente.

Si consideri un'onda divergente $\varphi(r,t)=f(ct-r)/r$ e si supponga che, all'istante t_1, l'oscillazione (rappresentata dalla funzione φ), sia compresa tra due sfere di raggio r_1 e $r_1+\varepsilon$ (cfr. fig. 7), il che significa che, in quest'istante, la funzione f è nulla al di fuori dello spazio compreso tra queste sfere, cioè $f(u)=0$ se $u<ct_1-r_1-\varepsilon$, oppure $u>ct_1-r_1$. Se ne deduce che in un istante successivo t_2 l'oscillazione è compresa tra due sfere di raggio $r_2=r_1+c(t_2-t_1)$ e $r_2+\varepsilon$.

Si può quindi concludere che il fronte d'onda (che separa la regione in quiete da quella interessata dall'oscillazione) si propaga con velocità c e che la regione interessata dall'oscillazione mantiene uno spessore costante ε. Nel caso limite di una sorgente puntiforme O (o distribuita su una sfera di raggio $r=r_1$) all'istante t_1, l'oscillazione all'istante t è limitata al fronte d'onda, cioè alla sfera di raggio $r_1+c(t-t_1)$. Questo fenomeno, che non è descrivibile dall'equazione delle onde a due dimensioni (o a un numero pari di dimensioni), è detto principio di Huygens, ed è fondamentale per la percezione distinta dei segnali sonori (o luminosi: cfr. oltre).

Si noti che, in questo caso, l'ampiezza φ si propaga con una riduzione inversamente proporzionale alla distanza: le curve che rappresentano in ogni punto dello spazio la variazione di φ in funzione del tempo, hanno tutte la stessa forma: esse differiscono solo per una traslazione e per un cambiamento di scala.

3.2. Fenomeni legati al moto della sorgente. Bang supersonico.

Si consideri una sorgente puntiforme di oscillazioni che si sposti nel tempo, ad esempio un aereo in volo. Da ogni punto della traiettoria parte, nell'istante in cui vi si trova l'aereo, una perturbazione del mezzo circostante. Ad un istante t, in cui l'aereo si trova nel punto M, le oscillazioni t_1, t_2, t_3 si trovano sulle sfere S_1, S_2, S_3, ..., di centro M_1, M_2, M_3, ... e raggio $c(t-t_1)$, $c(t-t_2)$, $c(t-t_3)$, ..., le cui rispettive posizioni dipendono dalla velocità v della sorgente (l'aereo) rispetto alla velocità di propagazione c (velocità del suono). Si supponga, dapprima, che la velocità dell'aereo sia inferiore alla velocità c del suono nell'aria: allora

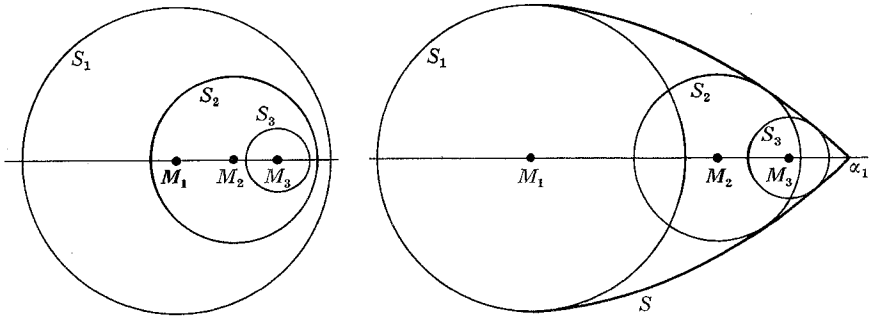


Figure 8-9.

Nel moto subsonico un osservatore al suolo riceve le onde sonore nell'ordine in cui sono emesse (fig. 8) mentre nel moto supersonico l'osservatore riceve contemporaneamente i rumori emessi da infiniti punti vicini (fig. 9). La rottura del muro del suono crea un'onda balistica che raggiunge l'osservatore se l'aereo vola basso.

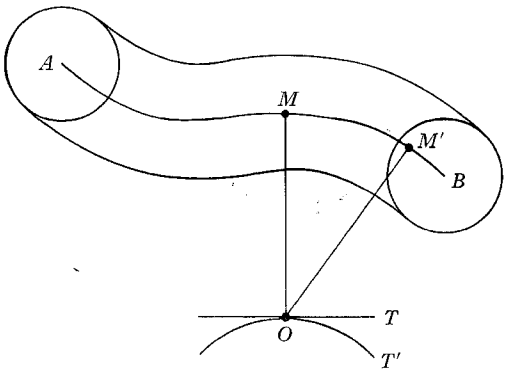


Figura 10.

La spiegazione del tuono con la teoria dell'onda balistica. Il lampo mette in contemporanea oscillazione tutti i punti della traiettoria da A a B: il rombo del tuono corrisponde all'arrivo di fronti d'onda successivi.

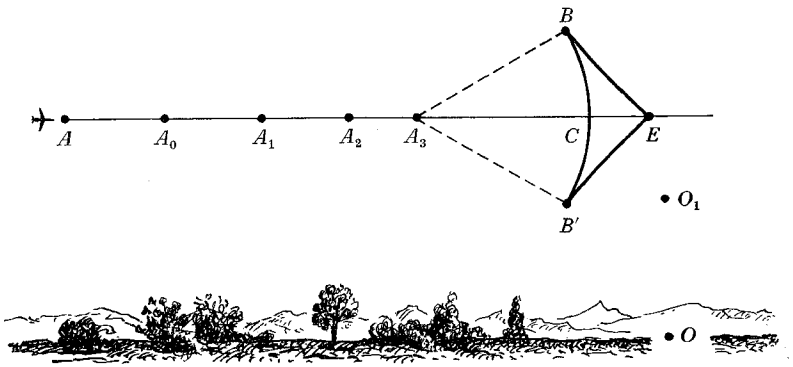


Figura 11.

Onda balistica provocata da un aereo che in A_1 raggiunge la velocità del suono e la supera. Se l'aereo vola basso, l'osservatore O percepisce due scoppi violenti.

la lunghezza di un arco di traiettoria M_1M_2 è sempre inferiore alla differenza dei raggi delle due sfere corrispondenti S_1 e S_2: le differenti sfere sono, in ogni istante t, interne l'una all'altra (fig. 8). Queste, durante la loro propagazione, arrivano a un osservatore al suolo nell'ordine in cui sono state emesse, generando la sensazione di un rumore continuo che si affievolisce con la distanza.

Se invece la velocità dell'aereo è supersonica, in un istante t due sfere S_1 e S_2 relative a due punti M_1 e M_2 della traiettoria s'intersecano, e l'insieme di queste sfere ammette un inviluppo S (fig. 9). Su un punto della superficie S arrivano contemporaneamente, all'istante t, i rumori emessi da una serie di punti vicini della traiettoria: questo è un punto di accumulo di rumore, che viene definito onda balistica o onda di Mach.

La teoria dell'onda balistica si applica senz'altro alla spiegazione del tuono. Si può ammettere, infatti, che il lampo sia un fenomeno istantaneo, dovuto alla contemporanea oscillazione di tutti i punti della sua traiettoria AB (fig. 10); l'onda balistica in un istante t è l'inviluppo, tubolare, dei fronti d'onda, sfere di raggio identico, centrate sulla traiettoria. Il tuono corrisponde all'arrivo all'osservatore di un'onda balistica: se ne possono percepire molte (ai tempi T e T' per la figura 10, in cui: $OM = cT$, $OM' = cT'$); il rombo del tuono corrisponde, invece, all'arrivo di fronti d'onda successivi.

Nel caso di traiettoria rettilinea della sorgente, a velocità supersonica v costante, la zona di silenzio (punti non interessati dal rumore) è, in ogni istante, l'esterno di un cono, detto cono di Mach, il cui semiangolo al vertice α è dato dalla relazione $\sin\alpha = c/v = 1/\mathfrak{M}$ ($\mathfrak{M}$ è detto numero di Mach).

Nel caso di un aereo che, aumentando la velocità, «sfondi il muro del suono» (cioè raggiunga e superi la velocità del suono), si ottiene un'onda balistica la cui sezione longitudinale è una curva chiusa da due segmenti (fig. 11): l'aereo ha velocità subsonica nell'intervallo A_0A_1, raggiunge la velocità del suono c nel punto A_1, raggiunge e conserva, a partire dal punto A_2 superato al tempo t_2, la velocità supersonica v. L'arco BCB' dell'onda balistica all'istante t è l'inviluppo dei fronti d'onda delle oscillazioni generate dai punti dell'arco AA_3, mentre la spezzata BEB' corrisponde alle oscillazioni originate dai punti dell'arco A_3E, essendo E l'estremità della traiettoria dell'aereo all'istante t.

Se l'aereo è sufficientemente alto rispetto all'osservatore (posizione O della figura), questi non è raggiunto dall'onda balistica e sente solo il rumore continuo dei reattori; se, invece, l'aereo vola basso, l'osservatore (posizione O_1) è raggiunto due volte dalle oscillazioni e sente due scoppi, la cui violenza è molto grande rispetto ai rumori di scappamento. La sovrapressione che corrisponde all'onda balistica di un aereo in volo supersonico può arrivare a decine di chilogrammi per metro quadrato.

4. *Caso generale di propagazione governata dall'equazione delle onde.*

I problemi di meccanica dei mezzi continui si risolvono spesso con equazioni alle derivate parziali, alle quali è necessario aggiungere delle condizioni ai limiti o delle condizioni iniziali. Attratti dalla semplicità e dal rigore della formulazione

di questi problemi, i grandi personaggi della matematica della metà del XIX secolo (Lagrange, Laplace, Poisson, Gauss, Cauchy) si sono applicati alla loro risoluzione. I loro sforzi hanno dato un impulso considerevole all'analisi matematica.

Il problema fondamentale per lo studio della propagazione è quello dei valori iniziali – detto anche problema di Cauchy in quanto, fra i considerevoli lavori di questo matematico, si trova la soluzione generale dei sistemi detti oggi «ben posti nel senso di Cauchy», e delle funzioni analitiche. Nel caso dell'equazione delle onde (cfr. § 3, n. 5) il problema di Cauchy consiste nella ricerca di una soluzione φ di questa equazione, di cui si conoscono, in un istante iniziale, ad esempio $t=0$, sia il suo valore sia quello della sua derivata rispetto al tempo: $\partial\varphi/\partial t$. Nel caso in cui l'incognita φ sia la pressione p di un gas (esempio dell'acustica), si è visto che $\partial p/\partial t$ è nota, quando è noto il vettore velocità $\mathbf{u}$: il problema consiste, dunque, noto lo stato del gas, considerato illimitato in un dato istante (pressione e vettore velocità), nel determinare il suo stato ulteriore. Il problema matematico è quello della soluzione dell'equazione delle onde con i «dati iniziali» per $t=0$ (si considera il punto dello spazio X di coordinate: x_1, x_2, x_3)

$$\varphi(X,0)=\varphi_0(X) \qquad \frac{\partial\varphi}{\partial t}(X,0)=\varphi_1(X).$$

È noto, per il precedente studio sulle onde sferiche, che la funzione

$$\psi(X,t)=\frac{f(|X-\xi|-ct)}{|X-\xi|}$$

(dove ξ è un punto fisso arbitrario e $|X-\xi|$ è la distanza dei punti X e ξ) è una soluzione dell'equazione delle onde. Poiché questa è lineare, è naturale pensare che la soluzione corrispondente ai dati iniziali sia una sovrapposizione di onde sferiche, centrate in differenti punti ξ, di ampiezza $F(\xi)$ che dipende dai dati iniziali in quel punto. Si scriverà

$$\varphi(X,t)=\int_{\mathbf{R}^3} F(\xi)\frac{f(|X-\xi|-ct)}{|X-\xi|}\,d\xi$$

dove $d\xi=d\xi_1 d\xi_2 d\xi_3$ indica l'elemento di volume dello spazio.

Se come f si sceglie la δ di Dirac, si ottiene

$$\varphi(X,t)=\frac{1}{ct}\int_{S_{ct}} F\,dS$$

dove S_{ct} è la sfera di raggio ct e centro X e dS è il suo elemento di superficie. In coordinate polari, l'integrale precedente si scrive

$$\varphi(X,t)=ct\int_0^{\pi}\int_0^{2\pi} F(X+ct\boldsymbol{\nu})\sin\alpha\,d\alpha\,d\beta$$

dove $\boldsymbol{\nu}$ è il vettore unitario nella direzione da X a ξ, cioè

$$\boldsymbol{\nu}=(\sin\alpha\cos\beta,\ \sin\alpha\sin\beta,\ \cos\alpha).$$

Si vede che la funzione φ, ottenuta scegliendo f come indicato, si annulla per $t=0$, mentre

$$\frac{\partial\varphi}{\partial t}(X,0)=4\pi c F(X).$$

D'altra parte si nota che la derivata parziale $\partial\varphi/\partial t$ di una soluzione φ dell'equazione delle onde soddisfa anche questa equazione e che quindi $(\partial^2\varphi/\partial t^2)=c^2\Delta\varphi$ è nulla per $t=0$, se questo vale per φ.

Si dimostra, infine, che la soluzione dell'equazione delle onde che soddisfa le condizioni iniziali proposte è data dalla formula di Poisson

$$\varphi(X,t)=tM_{ct}(\varphi_1)+\frac{\partial}{\partial t}\{tM_{ct}(\varphi_0)\}$$

dove M_{ct} indica la media relativa alla sfera di raggio ct e di centro X, per esempio

$$M_{ct}(\varphi_1)=\frac{1}{4\pi}\int_0^{\pi}\int_0^{2\pi}\varphi_1(x_1+ct\sin\alpha\cos\beta,\ x_2+ct\sin\alpha\sin\beta,\ x_3+ct\cos\alpha)\sin\alpha\,d\alpha\,d\beta.$$

Nella formula di Poisson si vede che la funzione φ (la pressione p nel caso dell'esempio dell'acustica), in un punto X all'istante t, dipende solo dai dati iniziali (lo stato iniziale nell'esempio della dinamica dei gas) nei punti che sono su una sfera di raggio ct e di centro X. Se i valori non nulli di $\varphi_0(X)$ e $\varphi_1(X)$ (l'oscillazione iniziale al tempo $t=0$) sono limitati in una regione D, allora in un punto esterno a D la soluzione è nulla (il gas resta in quiete) fino a che ct non sia uguale alla minima distanza da X a D. Essendo questa distanza su una normale da X alla frontiera di D, il fronte d'onda all'istante t (che separa i punti in-

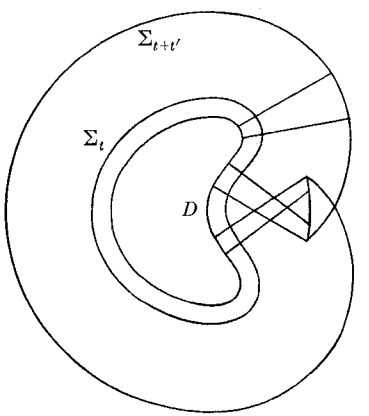


Figura 12.
Ripiegamento del fronte d'onda.

teressati da quelli non interessati dall'oscillazione) è ottenuto aggiungendo la lunghezza ct a tutte le normali alla superficie frontiera di D. I fronti d'onda successivi (detti anche superfici d'onda) hanno le stesse normali, che si definiscono raggi. La velocità di propagazione di un fronte d'onda, nella direzione di un raggio, è uguale a c.

Se la superficie frontiera di D è concava (fig. 12), il fronte d'onda, dopo un certo tempo, si ripiega su se stesso: i raggi ammettono un inviluppo, detto caustica, in cui vi è un accumulo di energia (si veda piú avanti l'analogo fenomeno ottico).

Il principio di Huygens, menzionato nel caso delle onde sferiche, è valido anche in questo caso: il valore di φ, nel punto X e all'istante t, dipende solo dai dati iniziali *sulla* sfera di centro X e raggio ct. Se il raggio di questa sfera è sufficientemente grande da contenere nel suo interno l'intera regione D, φ è di nuovo nullo nel punto X: nello spazio tridimensionale, un'oscillazione iniziale confinata a una regione limitata genera un'oscillazione che dura un tempo finito per ogni osservatore.

5. *Propagazione di moti periodici (piccole oscillazioni, vibrazioni).*

I fenomeni periodici svolgono un ruolo essenziale nell'osservazione della natura. I suoni sono prodotti e trasmessi da vibrazioni, cioè da oscillazioni periodiche di corpi materiali; anche la luce e le onde radio derivano dalla propagazione di fenomeni periodici.

Fin dal XVIII secolo questa realtà fisica ha portato i matematici a considerare somme di funzioni trigonometriche per risolvere i corrispondenti problemi. Fourier, in una memoria fondamentale letta il 21 dicembre 1807 all'Académie des Sciences di Parigi, mostrò tutta la potenza di questo metodo, mentre Poisson e Dirichlet fornirono rigorose prove della convergenza delle serie di Fourier di funzioni ampiamente arbitrarie (dette a variazione limitata); questa convergenza, presa in un senso piú generale, si applica, ora, anche alle funzioni generalizzate, o distribuzioni di Schwartz.

Ogni funzione periodica f, di periodo T – cioè di frequenza $\omega = 2\pi/T$ – è scritta come somma di funzioni sinusoidali di periodo T, $T/2$, $T/3$, ... e, dunque, di frequenze ω, 2ω, 3ω, ...:

$$f(t) = a_1 \cos(\omega t - \varphi_1) + a_2 \cos(2\omega t - \varphi_2) + \dots$$

Nel caso in cui si applichi il principio di sovrapposizione (acustico, ottico, ecc.), per studiare i movimenti periodici è sufficiente studiare i movimenti sinusoidali caratterizzati da una funzione (o un vettore): $u = a \cos(\omega t - \varphi)$. La frequenza ω è un numero indipendente dal punto X dello spazio, l'ampiezza a e la fase φ sono delle funzioni di questo punto, indipendenti dal tempo t.

5.1. Propagazione di un'onda piana sinusoidale.

Un'applicazione dei risultati generali enunciati nei paragrafi precedenti per le oscillazioni di una corda o per le onde piane in un fluido, che soddisfino l'equazione di d'Alembert, dimostra che se l'oscillazione del punto $x = 0$ è l'oscillazione sinusoidale $u_0 = a \cos \omega t$, la corrispondente onda progressiva è $u = a \cos \omega [t - (x/c)]$. Si definiscono lunghezza d'onda λ e numero d'onda k i numeri $\lambda = c(2\pi/\omega) = cT$ e $k = 2\pi/\lambda$. L'onda progressiva è $u = a \cos(\omega t - kx)$ con $\omega/k = c$, oppure ancora è $u = a \cos kx'$ con $x' = x - ct$. In qualsiasi istante t, un'onda piana progressiva è una sinusoide di periodo λ nello spazio. Quando t varia, questa sinusoide si muove, nella direzione ox, con un movimento uniforme, di velocità $\omega/k = c$.

Generalmente si definisce velocità di fase di un'onda piana sinusoidale $u = a \cos(\omega t - kx)$ il rapporto $v = \omega/k$ che esprime la velocità di spostamento dei massimi e dei minimi. Nel caso in cui u soddisfi l'equazione di d'Alembert, la velocità di fase è uguale alla velocità c di propagazione dei fronti d'onda.

Nel caso di un'onda piana sinusoidale progressiva in un fluido perfetto, se la velocità u è data da $u = a \cos(\omega t - kx)$, la perturbazione della pressione è

$$\frac{\partial p}{\partial t} = -\rho_0 c^2 \frac{\partial u}{\partial x} = -\rho_0 c^2 k a \sin(\omega t - kx).$$

Dunque,

$$p = \rho_0 c^2 \frac{k}{\omega} a \cos(\omega t - kx),$$

si vede che velocità e pressione variano in fase, con ampiezze proporzionali. La velocità u agisce nella direzione della velocità di propagazione (cioè positiva per un'onda progressiva) per le perturbazioni di pressione positive (compressioni); nella direzione contraria per le negative (rarefazioni).

6. *Propagazione delle onde luminose.*

La natura ondulatoria o corpuscolare della luce ha dato luogo ad una lunga polemica non ancora conclusasi ai giorni nostri. Fin dal 1673 fu abbozzata da Huygens una teoria ondulatoria, partendo dalla sua conoscenza dell'esperienza di Bontholin (1670), sulla doppia rifrazione dello spato d'Islanda; ciò nonostante, si preferí la teoria corpuscolare – anche dopo Newton – fino alla scoperta delle interferenze fatta da Young nel 1807, e soprattutto fino alle esperienze di Fresnel sull'interferenza e la diffrazione e l'interpretazione teorica da parte di Fresnel e Arago nel 1815 e 1819 che portarono definitivamente a considerare la luce come una grandezza a variazione sinusoidale (anche se la nozione di onde luminose deve essere rivista nella nuova meccanica quantistica, in dualismo con l'esperienza del fotone).

I fenomeni di polarizzazione, scoperti nel 1807 da Malus (che credeva nella

teoria corpuscolare), confermati e formalizzati da Fresnel entro la teoria ondulatoria, portano a concludere che le onde luminose sono onde trasversali.

L'ammirevole sintesi di Maxwell, che identifica le onde luminose con particolari onde elettromagnetiche, categoria di cui fanno parte anche le onde hertziane, le onde infrarosse e i raggi X, ha ricevuto una clamorosa conferma dalle esperienze di Hertz nel 1888. Le equazioni di Maxwell sono una pietra miliare nella fisica contemporanea. La loro apparente contraddizione con la meccanica newtoniana (sono invarianti per il gruppo di Lorentz ma non per il gruppo di Galileo) fu all'origine di quella rivoluzione concettuale costituita dall'avvento della relatività ristretta proposta da Einstein nel 1905.

In un dielettrico omogeneo perfetto, le equazioni di Maxwell che governano l'evoluzione dei campi elettrico **E** e magnetico **H**, sono

$$\operatorname{div}\mathbf{E}=0 \qquad \operatorname{div}\mathbf{H}=0 \qquad \frac{\varepsilon}{c}\frac{\partial\mathbf{E}}{\partial t}=\operatorname{rot}\mathbf{H} \qquad \frac{\mu}{c}\frac{\partial\mathbf{H}}{\partial t}=-\operatorname{rot}\mathbf{E}$$

dove ε e μ sono costanti, uguali a 1 nel caso del vuoto, c è il rapporto delle unità elettromagnetiche ed elettrostatiche di quantità di elettricità.

Dalle equazioni di Maxwell, riportate sopra, si deduce che il vettore **E** soddisfa la seguente equazione delle onde:

$$\frac{\partial^2\mathbf{E}}{\partial t^2}=v^2\Delta\mathbf{E} \qquad \text{con } v^2=\frac{c^2}{\varepsilon\mu}.$$

La velocità di propagazione di un fronte d'onda luminoso è, dunque, uguale a v (a c nel vuoto).

Nei mezzi trasparenti reali le equazioni che governano l'evoluzione dei campi sono piú complesse di quelle qui riportate – che sono valide solo con un'approssimazione piú o meno buona. Si può osservare che (tranne nel vuoto) la velocità di propagazione di un'onda luminosa dipende dalla sua frequenza, poiché l'indice di rifrazione n (legato alla velocità v dalla formula $n=c/v$) dipende dal colore della luce incidente. Si dice che c'è dispersione di onde (si veda oltre).

6.1. Polarizzazione.

Siano E_1, E_2, E_3 le tre componenti del vettore campo elettrico **E** rispetto agli assi cartesiani $oxyz$. Ogni componente soddisfa l'equazione delle onde, per esempio $\partial^2E_1/\partial t^2=(v^2)\Delta E_1$. Un'onda piana sinusoidale, che si propaghi nella direzione ox e sia soluzione dell'equazione delle onde, è dunque tale che

$$E_1=A_1\cos\left[\omega\left(t-\frac{x}{v}\right)-\varphi_1\right]$$
$$E_2=A_2\cos\left[\omega\left(t-\frac{x}{v}\right)-\varphi_2\right]$$
$$E_3=A_3\cos\left[\omega\left(t-\frac{x}{v}\right)-\varphi_3\right],$$

dove A_1, A_2, A_3, φ_1, φ_2, φ_3 sono costanti.

D'altra parte, dalla condizione

$$\operatorname{div}\mathbf{E}=\frac{\partial E_1}{\partial x}+\frac{\partial E_2}{\partial y}+\frac{\partial E_3}{\partial z}=0$$

si deduce che $A_3=0$, dunque $E_3=0$; il vettore **E** è parallelo al piano d'onda oyz: la vibrazione è trasversale. Il vettore di componenti E_1, E_2, descrive, nel piano d'onda, un'ellisse.

Nel caso in cui il vettore **E** abbia una direzione fissa, si dice che l'onda ha polarizzazione rettilinea. Prendendo oy nella direzione di **E**, si ha

$$E_1=0 \qquad E_2=A\cos\left[\omega\left(t-\frac{x}{v}\right)-\varphi\right] \qquad E_3=0.$$

Il campo magnetico associato è

$$H_1=0 \qquad H_2=0 \qquad H_3=\sqrt{\frac{\varepsilon}{\mu}}A\cos\left[\omega\left(t-\frac{x}{v}\right)-\varphi\right]$$

cioè è ortogonale, ma in fase con il vettore campo elettrico.

La luce naturale monocromatica (proveniente dal Sole, da un arco elettrico, ecc.) è costituita da vibrazioni ellittiche, di cui forma, orientamento e fase cambiano in modo irregolare piú di un milione di volte al secondo; la luce riflessa è parzialmente polarizzata, in dipendenza dall'ampiezza dell'angolo d'incidenza.

6.2. Energia trasportata dall'irraggiamento elettromagnetico.

La densità di energia di un campo elettromagnetico è, nelle unità correnti, $U=(1/8\pi)(\varepsilon\mathbf{E}^2+\mu\mathbf{H}^2)$, cioè, per l'onda piana polarizzata,

$$U=\frac{1}{4\pi}\varepsilon A^2\cos^2\left[\omega\left(t-\frac{x}{v}\right)-\varphi\right].$$

Anche quest'energia si propaga con velocità v, nella direzione ox perpendicolare ai piani d'onda (direzione dei raggi luminosi corrispondenti all'onda considerata).

6.3. Ottica geometrica.

L'ottica geometrica si basa sulla nozione di raggio luminoso: si definiscono cosí le traiettorie ortogonali delle superfici (o fronti) d'onda, che sono rette nel caso dell'equazione delle onde, equazione a coefficienti costanti che governa la propagazione del campo elettromagnetico (e in particolare della luce) in un ambiente omogeneo e isotropo. La «legge di propagazione rettilinea della luce» in tale ambiente, insieme alle leggi di riflessione e di rifrazione attraverso la superficie di separazione di due mezzi con indice differente, costituiscono i fondamenti dell'ottica geometrica. Se ne possono enunciare le leggi senza ricorrere alla natura ondulatoria (né corpuscolare) della luce: queste hanno ispirato in modo particolare le ricerche di un grande nome della meccanica analitica, Hamilton, che pose, nel 1833, le basi del calcolo variazionale nel tentativo di riordinare dei risultati suscettibili d'interpretazione ondulatoria o corpuscolare per mezzo del

principio di minima azione di Fermat, Maupertuis o Lagrange. Le leggi dell'ottica geometrica possono essere verificate sperimentalmente, con un'approssimazione limitata dai fenomeni di diffrazione dovuti alla natura ondulatoria della luce, che rendono impossibile l'isolamento fisico di un «raggio luminoso», limitandolo con diaframmi di piccolissima apertura. Tuttavia, i matematici moderni hanno dimostrato, a partire dalla teoria ondulatoria, che le leggi dell'ottica geometrica hanno una validità asintotica e approssimata e sono tanto piú verificate quanto maggiore è la frequenza della radiazione luminosa. I metodi di sviluppo asintotico, nati in questo modo, hanno trovato un largo campo di applicazione sia in fisica sia in matematica pura (Luneburg, Lax, Leray). Piú avanti si esporrà il principio del calcolo.

Il campo elettrico corrispondente a un fascio di luce monocromatica di frequenza ω è rappresentato da un vettore della forma

$$E_1 = a_1 \cos\{\omega[t-\sigma(x)]-\varphi_1(x)\}$$
$$E_2 = a_2 \cos\{\omega[t-\sigma(x)]-\varphi_2(x)\}$$
$$E_3 = a_3 \cos\{\omega[t-\sigma(x)]-\varphi_3(x)\},$$

dove a_1, a_2, a_3, come φ_1, φ_2, φ_3, sono funzioni a priori arbitrarie del punto X dello spazio, ma si suppone che queste varino poco (rispetto a ω) da un punto all'altro. Scrivendo che $\mathbf{E}$ deve verificare l'equazione delle onde $\sum_{i=1}^{3}(\partial^2\mathbf{E}/\partial x_i^2)-(1/v^2)(\partial^2\mathbf{E}/\partial t^2)=0$, a meno di termini di ordine ω, si trova che la soluzione deve verificare l'equazione del primo ordine alle derivate parziali (detta eiconale)

$$[\operatorname{grad}\sigma]^2 \equiv \sum_{i=1}^{3}\left(\frac{\partial\sigma}{\partial x_i}\right)^2 = v^2.$$

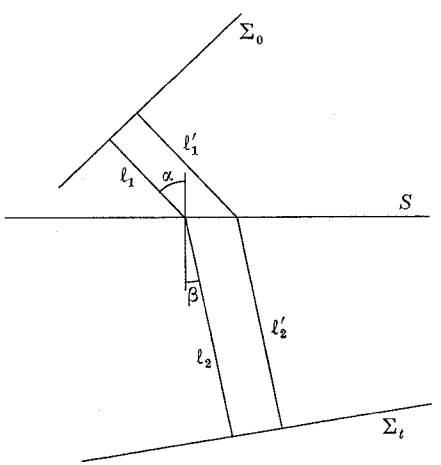


Figura 13.
Leggi della riflessione e della rifrazione.

Facendo un paragone con le definizioni già date, si vede che le superfici $\sigma(x)=t$ sono i fronti d'onda all'istante t, mentre le traiettorie ortogonali di questi fronti d'onda (tangenti in ogni punto al vettore $\operatorname{grad}\sigma$) sono dette raggi.

Scrivendo che l'equazione delle onde è verificata a meno di termini di ordine 1, si trova che le fasi φ_1, φ_2, φ_3 si propagano senza modifiche lungo i raggi; soddisfano, ad esempio, l'equazione

$$\operatorname{grad}\sigma\cdot\operatorname{grad}\varphi_1 = \sum_{i=1}^{3}\frac{\partial\sigma}{\partial x_i}\frac{\partial\varphi_1}{\partial x_i} = 0$$

mentre le ampiezze parziali a_1, a_2, a_3 si propagano lungo i raggi con uno smorzamento (o amplificazione) che dipende dalla divergenza (o convergenza) di questi ultimi; ad esempio $\operatorname{grad}\sigma\cdot\operatorname{grad}a_1+(1/2)a_1\Delta\sigma=0$, che si riduce per le onde piane (raggi paralleli) all'equazione di conservazione lungo i raggi $\operatorname{grad}\sigma\cdot\operatorname{grad}a_1=0$. Dall'ortogonalità fra i raggi e le superfici d'onda e dalla costruzione di queste, si deducono le leggi della riflessione e della rifrazione (cfr. fig. 13). Se si indica con S la superficie di separazione dei mezzi con indice n_1 e n_2 (velocità c/v_1 e c/v_2), e con Σ_0 e Σ_t le superfici d'onda all'istante zero e all'istante t, si ha $t=t_1+t_2=t_1'+t_2'$; $\ell_1=v_1t_1$; $\ell_1'=v_1t_1'$; $\ell_2=v_2t_2$; $\ell_2'=v_2t_2$, da cui

$$\frac{n_1}{n_2}=\frac{v_2}{v_1}=\frac{\sin\alpha}{\sin\beta}.$$

7. *Equazioni iperboliche.*

Si sono studiati, fino ad ora, i piú semplici fenomeni di propagazione, cioè quelli governati dall'equazione delle onde.

Un certo numero di fenomeni classici, di ampiezza troppo piccola perché i loro effetti siano additivi, sono governati da equazioni alle derivate parziali, che sono, come l'equazione delle onde, equazioni di tipo iperbolico: una soluzione che, in un dato istante t_0, è nulla al di fuori di una regione limitata D_0 è, in un successivo istante t, nulla al di fuori di un'altra regione limitata D_t, che dipende da D_0 e dall'equazione: in altre parole, al fenomeno è associato un fronte d'onda che si propaga con una velocità finita dipendente dall'equazione e, eventualmente, anche dalla direzione di propagazione.

L'iperbolicità di un'equazione alle derivate parziali dipende solo dai suoi termini principali, cioè dalla derivata di ordine piú alto. Un'equazione del secondo ordine è iperbolica, se la forma quadratica definita dai coefficienti delle derivate seconde dell'incognita è di tipo iperbolico (non degenere e con tutti i termini quadratici di una decomposizione canonica dello stesso segno salvo uno). L'equazione di d'Alembert, l'equazione delle onde sono iperboliche. L'equazione di Webster, in una variabile spaziale, che governa la propagazione del suono con onde piane di un altoparlante

$$\frac{\partial^2 u}{\partial t^2}-v^2\left(\frac{\partial^2 u}{\partial x^2}+\alpha\frac{\partial u}{\partial x}\right)=0$$

è anch'essa un'equazione iperbolica: si dimostra che una soluzione nulla al di fuori dell'intervallo $a \leq x \leq b$ all'istante zero, è nulla all'istante t, al di fuori dell'intervallo $a - vt \leq x \leq b + vt$, dunque i fronti d'onda si propagano con velocità v, come nel caso dell'equazione di d'Alembert. Un risultato analogo è valido anche per l'equazione seguente, che ha gli stessi termini principali dell'equazione delle onde:

$$\frac{\partial^2 u}{\partial t^2} - v^2 \Delta u + \sum_{i=1}^{3} a_i \frac{\partial u}{\partial x_i} + bu + c = 0;$$

un suo caso particolare è l'equazione di Klein-Gordon, soddisfatta dalle onde di Broglie delle particelle libere relativistiche, di massa m:

$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} - \Delta u + mu = 0.$$

Il principio di Huygens non è valido, salvo eccezioni, per queste equazioni, e neanche per equazioni piú generali in cui i coefficienti non sono costanti: dopo il passaggio del fronte d'onda, le oscillazioni persistono. Si dice che i casi eccezionali possiedono delle «lacune». Tutti i problemi ben posti di evoluzione della dinamica relativistica dei mezzi continui – e molti di quelli della meccanica classica – si risolvono con equazioni iperboliche alle derivate parziali. Sono stati ottenuti risultati molto generali da Leray nel 1952.

7.1. Dispersione.

Anche nel caso in cui i fenomeni sono governati da un'equazione iperbolica, lo studio delle soluzioni particolari periodiche, in particolare sinusoidali, ha spesso destato interesse: queste soluzioni corrispondono per lo piú a «vibrazioni forzate», regime indotto da una sorgente di oscillazioni di frequenza fissa ω; lo studio della loro propagazione è molto complesso.

Un'onda piana sinusoidale di frequenza ω e numero d'onda k, che si propaghi nella direzione ox, si scrive $u(x,t) = a(x) \cos(\omega t - kx)$. La velocità di fase $V = \omega/k$ caratterizza gli spostamenti nel tempo dei massimi e dei minimi successivi dell'onda – non è una velocità di segnale (si veda oltre). Se, per differenti onde di una stessa classe di fenomeni – soluzioni o no di equazioni iperboliche –, la velocità di fase è diversa, si dice che si hanno delle onde dispersive. La relazione che lega k e ω è detta relazione di dispersione. Se questa relazione implica che k sia immaginario per alcuni valori di ω, questi valori non corrispondono a nessuna onda sinusoidale della famiglia e si dice che queste sono al di fuori della banda passante di frequenza.

Esempio.

Si cerchino le onde sinusoidali che sono soluzione dell'equazione di Webster menzionata prima per studiare la propagazione di un suono di frequenza data ω in un altoparlante.

Si consideri il caso in cui α è costante (diffusore esponenziale $S = S_0 e^{\alpha x}$). La funzione $u = a(x) \cos(\omega t - kx)$ è una soluzione dell'equazione

$$\frac{\partial^2 u}{\partial t^2} - v^2 \left(\frac{\partial^2 u}{\partial x^2} + \alpha \frac{\partial u}{\partial x} \right) = 0,$$

se

$$2 \frac{da}{dx} + \alpha a = 0 \qquad \frac{d^2 a}{dx^2} + \frac{da}{dx} \alpha + \left(\frac{\omega^2}{v^2} - k^2 \right) a = 0$$

cioè l'ampiezza a è $a = a_0 e^{-\alpha x/2}$. Quando la relazione di dispersione è $k^2 = (\omega^2/v^2) - (\alpha^2/4)$, si ottiene un valore di k solo se la frequenza ω è superiore al valore $\omega_0 = \alpha v/2$; in un diffusore esponenziale si può avere propagazione di una vibrazione sinusoidale (suono puro) solo se la sua frequenza è superiore a questa frequenza particolare ω_0, detta frequenza di taglio del diffusore.

La velocità di fase della vibrazione $u = a(x) \cos(\omega t - kx)$ è allora

$$\frac{\omega}{k} = \frac{\omega}{\sqrt{\frac{\omega^2}{v^2} - \frac{m^2}{4}}} = \frac{v}{\sqrt{1 - \frac{\omega_0^2}{\omega^2}}}.$$

Questa velocità di fase dipende dalla frequenza ω ed è sempre superiore alla velocità di propagazione v dei fronti d'onda. Qui la velocità di fase aumenta quando ω diminuisce, diventa infinita alla frequenza di taglio e ciò significa che, a questa frequenza, tutta l'aria dell'altoparlante vibra in fase, si ha $k = 0$, $u = a(x) \cos \omega t$. Si sottolinea che questo risultato non corrisponde in nessun modo a una velocità di propagazione infinita di un segnale sonoro: la partenza di un segnale (emissione di un suono all'entrata del diffusore) è seguita da un regime transitorio; il regime forzato di uguale frequenza si stabilisce solo in seguito, asintoticamente.

Si osservi che l'ampiezza dell'onda è proporzionale a $e^{-\alpha x/2}$; dunque, la potenza ricevuta per unità di superficie è proporzionale a $e^{-\alpha x}$ e, dal momento che la superficie è proporzionale a $e^{\alpha x}$, la potenza totale che attraversa una sezione è costante: la potenza acustica si trasmette completamente dalla sezione di entrata alla sezione di uscita del diffusore, per frequenze superiori alla frequenza di taglio.

Per frequenze inferiori a quella di taglio, la soluzione ha uno smorzamento esponenziale supplementare della forma

$$u = a e^{-[(\alpha/2) + \sqrt{(\alpha^2/4) - (\omega^2/v^2)}]\, x} \cos \omega t.$$

Si vede che, in pratica, l'altoparlante non trasmette i suoni di frequenza inferiore alla propria frequenza di taglio.

7.2. Treni d'onda, velocità di gruppo.

Quando la velocità di fase delle onde sinusoidali che sono soluzione di un problema dato dipende dalla loro frequenza, risulta un nuovo fenomeno: l'esi-

stenza di treni d'onda e una velocità di propagazione di questi treni d'onda, detta velocità di gruppo. La velocità di gruppo è stata introdotta piú dai fisici che dai matematici e svolge un ruolo importante rispetto all'esperienza. Questa riapparirà in meccanica quantistica.

Un esempio è fornito dall'analisi del fenomeno dei battimenti prodotti dalla sovrapposizione di due onde sinusoidali della stessa ampiezza a, di frequenza prossima ω_1 e ω_2, e di velocità di fase prossima V_1 e V_2 e, dunque, di numero d'onda k_1 e k_2 prossimo (si ricorda che $k=\omega/v$; e $\lambda=2\pi/k$ è la lunghezza d'onda). In un dato istante t, la sovrapposizione delle due onde è rappresentata nella figura 14. Un massimo dell'ampiezza risultante è ottenuto nei punti M, in cui coincidono dei massimi delle due onde componenti. Questa coincidenza non si mantiene nel tempo, in quanto le sinusoidi S_1 e S_2 si propagano con velocità diversa. Si supponga che S_2 sia la sinusoide con lunghezza d'onda piú grande e che si propaghi piú velocemente di S_1: il massimo M_2 che segue N_2, avrà raggiunto il massimo M_1 che seguiva N_1, per dare un nuovo massimo M di ampiezza risultante, in un tempo τ tale che $(V_2-V_1)\tau=\lambda_2-\lambda_1$. Questo massimo M è a una distanza $V_2\tau-\lambda_2$ da N.

La velocità media di propagazione del massimo dell'ampiezza risultante è dunque

$$U=V_2-\frac{\lambda_2}{\tau}=V_2-\lambda_2\frac{V_2-V_1}{\lambda_2-\lambda_1}$$

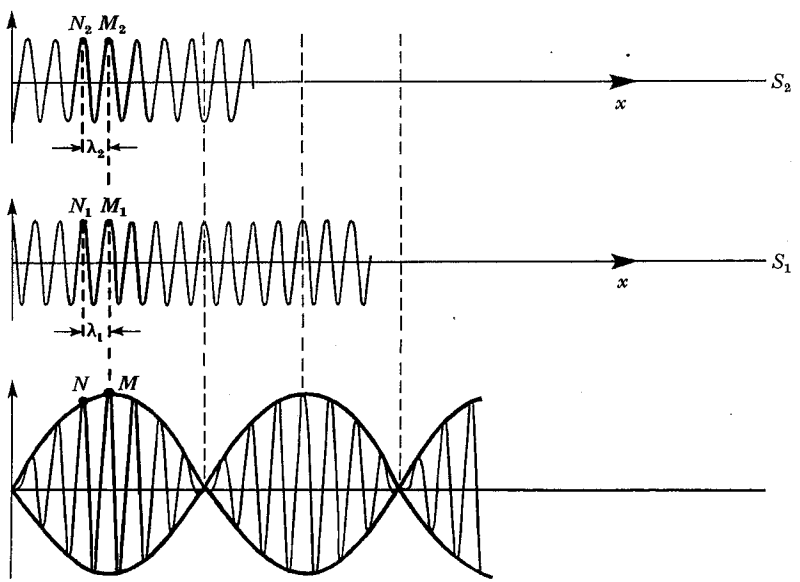


Figura 14.
Sovrapposizione delle onde S_1 e S_2.

da cui, considerando V_2-V_1 e $\lambda_2-\lambda_1$ come infinitamente piccoli, $U=V-\lambda(dV/d\lambda)$. Questi risultati si trovano immediatamente con il seguente calcolo della sovrapposizione delle due onde u_1 e u_2:

$$u=u_1+u_2=a\cos(\omega_1 t-k_1x)+a\cos(\omega_2 t-k_2x),$$

da cui

$$u=2a\cos\left(\frac{\omega_1+\omega_2}{2}t-\frac{k_1+k_2}{2}x\right)\cos\left(\frac{\omega_1-\omega_2}{2}t-\frac{k_1-k_2}{2}x\right).$$

Se si suppone che ω_1 e ω_2 da una parte, k_1 e k_2 dall'altra, siano prossimi rispettivamente a una frequenza ω e a un numero d'onda k dati, si ha

$$u\simeq 2a\cos(\omega t-kt)\cos\left(\frac{\omega_1-\omega_2}{2}t-\frac{k_1-k_2}{2}x\right),$$

da cui si deduce la rappresentazione grafica (fig. 14); la sinusoide S, inviluppo dei massimi delle oscillazioni di u in un dato istante t

$$u_S=2a\cos\left(\frac{\omega_1-\omega_2}{2}t-\frac{k_1-k_2}{2}x\right)$$

appare come separatrice dei gruppi, o treni d'onda. La sua velocità U di propagazione nel tempo è detta, per questa ragione, velocità di gruppo. Si ha $U=(\omega_1-\omega_2)/(k_1-k_2)$ da cui, se $\omega_1-\omega_2$ e k_1-k_2 sono considerati infinitamente piccoli, $U=d\omega/dk$, la quale, tenendo conto della relazione $k=2\pi/\lambda$, coincide con la formula scritta precedentemente.

Si vede che la velocità di gruppo U è inferiore alla velocità di fase V se questa è una funzione crescente di λ: le onde che costituiscono il gruppo sembrano, all'interno del gruppo, procedere in avanti. In caso contrario, si verifica il fenomeno inverso.

Esempio: quando si getta una pietra in uno stagno, si osserva un fenomeno del primo tipo, sebbene piú complesso; si produce un treno di onde circolari, nel cui interno le onde individuali si propagano piú velocemente di quelle dell'insieme: queste si smorzeranno nella parte anteriore del treno d'onda, mentre nella parte posteriore ne nascono altre. Lo studio matematico della propagazione delle onde superficiali di un liquido pesante (onda lunga) è in accordo con questa osservazione.

Si consideri una massa d'acqua in equilibrio nel campo gravitazionale; sia $oxyz$ una terna cartesiana, con oz verticale diretto verso l'alto, e il piano oxy coincidente con quello della superficie libera, orizzontale, del liquido a riposo. Si considerino, ora, delle onde rettilinee di piccola ampiezza parallele a oy che si propaghino nella direzione ox. Si indichino con ξ e ζ le componenti, secondo ox e oz, dello spostamento di un elemento del fluido all'istante t, che in equilibrio aveva come coordinate x e z. Si trova che questo spostamento corrisponde a un'onda sinusoidale che soddisfa le equazioni dell'idrodinamica dei fluidi pesanti (con una superficie libera), se $\xi=-ae^{kz}\cos(\omega t-kx)$, $\zeta=ae^{kz}\cos(\omega t-kx)$ a condizione che la profondità sia molto grande rispetto alla lunghezza d'onda $\lambda=2\pi/k$ e che la velocità di fase $V=\omega/k$ soddisfi la relazione $V=\sqrt{g\lambda/2\pi}$. Cioè

a condizione che $\omega = \sqrt{gk}$, dove con g viene indicata l'accelerazione di gravità.

La superficie libera all'istante t è il luogo dei punti degli elementi fluidi che, in equilibrio, appartenevano al piano $z = 0$. Le coordinate (X, Z) di un tale elemento, che dipendono dalla sua ascissa x in quiete, sono

$$X = x + \xi = x - a\cos(\omega t - kx)$$
$$Z = \zeta = a\sin(\omega t - kx).$$

Da queste relazioni si deduce che ogni elemento del fluido descrive, nel tempo, una piccola circonferenza di raggio a, attorno alla sua posizione di equilibrio, con uno sfasamento che si traduce in un profilo ondulatorio della superficie libera, ad ogni istante dato. Questo profilo si propaga con la velocità di fase V (fig. 15, in cui i profili sono stati disegnati in due istanti, che differiscono di $T/4$, con $T = 2\pi/\omega$).

L'altezza dell'onda lunga – distanza verticale fra una cresta e una gola – è uguale a $2a$. La forma del profilo dell'onda nelle vicinanze della sommità dipende dal rapporto $2a/\lambda$ di questa altezza rispetto alla lunghezza d'onda, ma, in questo caso, cessano di essere valide le approssimazioni fatte finora.

Dal momento che la velocità di fase è una funzione crescente della lunghezza d'onda λ, la velocità di gruppo del treno d'onda è inferiore alla velocità di fase delle onde individuali.

Nell'esempio precedente si è trascurato l'effetto della tensione superficiale del liquido. Se si tiene conto di questo effetto, si trova che la perturbazione verticale (in acqua profonda) è ancora della forma $Z = a\sin(\omega t - kx)$, a condizione che la velocità di fase $V = \omega/k$ soddisfi la relazione (in cui ρ indica la densità)

$$V^2 = \frac{g\lambda}{2\pi} + \frac{A}{\rho}\frac{2\pi}{\lambda} \qquad \text{cioè } \omega^2 = gk\left(1 + \frac{A}{\rho g}k^2\right).$$

Per l'acqua, in unità CGS, si ha $g = 981$, $\rho = 1$, $A = 74$ e, quindi, l'effetto della tensione superficiale diventa trascurabile (dell'ordine dell'1 per cento) per lunghezze d'onda inferiori a 10 cm; al contrario, questa svolge un ruolo molto importante nella propagazione di onde con lunghezza d'onda dell'ordine del centimetro, e un ruolo preponderante per lunghezze d'onda dell'ordine del millimetro (che si possono osservare, ad esempio, facendo vibrare con un diapason la superficie dell'acqua).

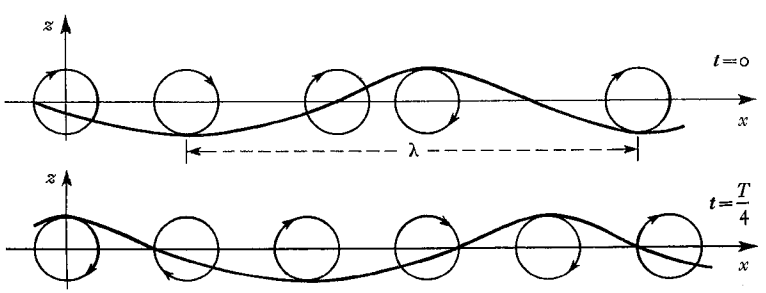


Figura 15.
Profilo ondulatorio della superficie libera in due istanti che differiscono di $\pi/2\omega$.

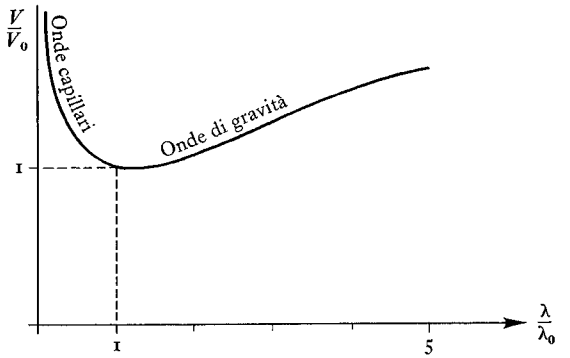


Figura 16.
Grafico delle velocità di fase e loro classificazione.

Per queste onde, dette capillari, la velocità di fase è una funzione decrescente della lunghezza d'onda e, dunque, inferiore alla velocità di gruppo.

Da un punto di vista del tutto generale, si vede che la velocità di fase è minima quando $(g\lambda/2\pi) = (A/\rho)(2\pi/\lambda)$, cioè $\lambda = \lambda_0 = 2\pi\sqrt{A/\rho g}$. Per ogni lunghezza d'onda $\lambda > \lambda_0$ ci sono due velocità di fase e dunque due frequenze possibili (cfr. fig. 16): una detta «capillare», e l'altra «di gravità».

Si vede anche che la velocità di gruppo $U(k) = \partial\omega/\partial k$ ammette un minimo U_m. Si trova, per l'acqua, $V_0 = 23{,}3$ cm/s, $\lambda_0 = 1{,}74$ cm e $U_m = 17{,}9$ cm/s.

7.3. Osservazione delle velocità di gruppo per una sorgente puntiforme istantanea.

Si consideri un treno d'onde generato da una perturbazione nel punto $r = 0$, all'istante $t = 0$. Si indichi con $U(k)$ la velocità di gruppo dell'onda di lunghezza $\lambda = 2\pi/k$; il treno d'onda corrispondente si troverà, all'istante t, alla distanza $r = U(k)t$. In altre parole, a una distanza r fissata si osserveranno dei treni d'onda la cui frequenza varia con il tempo secondo la legge $U(k) = r/t$. Per le onde di gravità, si ha

$$U(k) = \frac{\partial\omega}{\partial t} = \frac{1}{2}\sqrt{\frac{g}{k}}.$$

La relazione fornisce dunque $k = gt^2/4r^2$, $\omega = gt/2r$. Questa legge è stata verificata, con buona precisione, per onde generate da una tempesta nell'Oceano Pacifico, a una distanza di duemila miglia.

In scala minore, gettando una pietra in uno stagno, si verifica che, t secondi

piú tardi, la zona d'impatto è il centro di un disco tranquillo di circa $18t$ cm, corrispondente al minimo della velocità di gruppo. Oltre, ci sono due valori di k per ogni valore di r. Tuttavia, le lunghezze d'onda maggiormente eccitate all'istante iniziale (cioè dello stesso ordine di grandezza della pietra) hanno l'ampiezza piú grande: i cerchi principali saranno, all'incirca, di $r = U(\pi/\ell)t$ per una pietra di raggio ℓ.

7.4. Fase stazionaria.

Le onde che si sono considerate fino ad ora rappresentano solo schematicamente i fenomeni reali. I treni d'onda subiscono, propagandosi in un mezzo materiale, modificazioni piú complesse di quelle che si sono studiate: i fenomeni, anche di piccola ampiezza, non sono governati in modo esatto, e senza altre condizioni, da equazioni alle derivate parziali. Un mezzo materiale non è mai del tutto omogeneo e neanche illimitato. Tuttavia, possono essere ottenute interessanti informazioni sui moti che sono sovrapposizione di moti oscillatori, se si conosce la relazione di dispersione che essi soddisfano. Un metodo particolarmente valido è quello detto della fase stazionaria, che si utilizza per la valutazione di diverse soluzioni asintotiche, locali o globali. Quest'ultimo caso fornisce delle regole di quantificazione matematicamente rigorose e fisicamente intuitive (lavori di Leray-Maslov). Si dànno i principî del metodo valutando un integrale della forma

$$u(x,t) = \int_{-\infty}^{+\infty} f(k) \exp i(\omega t - kx)\, dk$$

che è la sovrapposizione di oscillazioni di frequenza ω, numero d'onda k e ampiezza $f(k)$. I numeri ω e k sono reali, e legati da una relazione di dispersione: ω è una funzione data di k. $f(k)$ è complessa, e contiene, a volte, l'ampiezza reale $|f(k)|$ e la fase $\varphi = \operatorname{Arg} f(k)$. Nei problemi di meccanica classica, l'onda fisica è la parte reale di u. Si ha

$$u(x,t) = \int_{-\infty}^{+\infty} f(k) \exp it\, \Phi(k, x/t)\, dx,$$

ponendo $\Phi(k, x/t) = \omega(k) - k(x/t)$ e interessa l'andamento della funzione u quando x e t sono grandi, piú precisamente quando nel rapporto x/t, di valore fissato U, t è grande.

È intuitivo, allora, il fatto che la funzione da integrare oscilli rapidamente con k, eccetto che vicino a quei valori di k per i quali la derivata di Φ rispetto a k si annulla: questi valori sono detti stazionari. L'integrale calcolato escludendo questi valori sarà molto piccolo (si dice che si integra su una regione «di interferenza distruttiva»). Il contributo principale all'integrale è dato dalle regioni «di interferenza costruttiva», cioè nelle vicinanze dei valori stazionari di k, per i quali

$$\frac{\partial \Phi}{\partial k}(k, U) = 0 \qquad \text{cioè } \omega'(k) = U.$$

Un calcolo matematico rigoroso fornisce la valutazione seguente, quando per i valori stazionari la derivata seconda $\omega''(k)$ è diversa da zero:

$$u(x,t) = \sum_{\substack{\text{valori stazionari } k,\\ \omega'(k)=x/t}} f(k) \sqrt{\frac{2\pi}{t\,|\omega''(k)|}} \exp i\left[\omega(k)t - kx + \frac{i\pi}{4} \operatorname{sign} \omega''(k) + o\left(\frac{1}{t^{3/2}}\right)\right].$$

Il fattore correttivo $o(1/t^{3/2})$ tende a zero t volte piú velocemente del termine principale. L'ampiezza di questo termine principale decresce come $1/\sqrt{t}$ (per un dato rapporto x/t). Questo risultato corrisponde alla diffusione spaziale proporzionale a $1/t$ della densità dell'energia (quadrato del modulo di u) corrispondente ai differenti numeri d'onda, poiché $x = t\omega'(k)$ comporta $dx = t\omega''(k)\, dk$.

In questo caso generale, si ritrova il risultato ottenuto piú sopra per le onde di gravità e che spesso viene formulato nel modo seguente: in una famiglia di onde dispersive, l'onda di numero k si propaga con la velocità di gruppo $U = \omega'(k)$. Il senso di questa frase è che l'osservatore posto a una distanza $x = Ut$ osserva soprattutto onde di frequenza k; si possono avere piú onde con questa frequenza, se queste hanno la stessa velocità di gruppo.

Si può sottolineare inoltre che il termine principale di $u(x,t)$ presenta uno sfasamento di $\pi/4$ rispetto all'onda di numero k.

Il caso in cui $\omega''(k) = 0$ per un valore di k tale che $\omega'(k) = 0$, richiede un trattamento particolare.

7.5. Velocità di segnale.

Si conviene in generale di definire segnale una perturbazione (elettrica, luminosa, acustica, ecc.) di un sistema a riposo, limitata nel tempo e nello spazio. La velocità di propagazione del segnale è, in linea di principio, determinata dal fronte dei punti interessati, in ogni istante dato. In pratica, questo fronte, e dunque la velocità apparente del segnale, dipende dalla sensibilità degli apparecchi che lo registrano. Cosí, anche per una perturbazione regolata da un'equazione iperbolica, in cui la velocità di propagazione del fronte è determinata dall'equazione, un apparecchio di sensibilità limitata può non rilevare le prime manifestazioni del segnale, molto deboli, ma solo le successive di maggiore intensità. Questo fenomeno avviene in particolare quando il segnale è costituito da una serie di oscillazioni e quando il sistema è dispersivo. La velocità di un segnale, allora, spesso coincide con la velocità di gruppo, ma non necessariamente. Questa è sempre inferiore alla velocità del fronte, punto importante nella teoria della relatività in cui nessun segnale può propagarsi con una velocità superiore a quella della luce nel vuoto.

Un esempio tipico del fenomeno precedente è quello della propagazione della luce in un mezzo materiale. Le proprietà ottiche di un corpo sono legate alle oscillazioni forzate, indotte dal campo elettrico dell'onda luminosa, degli ioni e degli elettroni che lo compongono. Il fronte d'onda di un segnale luminoso si propaga sempre con la velocità c della luce nel vuoto, ma l'intensità corrispondente può essere molto debole: questa intensità diventa apprezzabile solo quan-

do gli ioni e gli elettroni oscillano in modo coerente sotto l'azione del campo elettrico: la velocità del segnale (e dunque l'indice di rifrazione) corrispondente dipende dalla frequenza, cioè dal colore dell'onda incidente. La velocità del segnale si confonde con la velocità di gruppo, tranne nelle vicinanze di certe frequenze particolari (frequenze di assorbimento), che sono quelle delle oscillazioni libere degli ioni o elettroni nel mezzo; l'energia dell'onda luminosa è, allora, quasi assorbita dalla stabilizzazione di queste vibrazioni forzate (risonanza).

Una teoria soddisfacente – benché approssimata – è stata messa a punto da Brillouin e Sommerfeld. Grosso modo, la teoria è la seguente: le equazioni di Maxwell, che regolano la propagazione dei campi vettoriali elettrico **E** e magnetico **H**, sono (la permeabilità magnetica è sempre molto vicina all'unità)

$$\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t} = -\operatorname{rot}\mathbf{E} \qquad \frac{4\pi \mathbf{I}}{c} = \operatorname{rot}\mathbf{H}$$

dove **I** è la somma del flusso di spostamento $(1/4\pi)(\partial\mathbf{E}/\partial t)$ e di un flusso corrispondente allo spostamento reale delle cariche elettriche, supposte elettroni per semplicità (le cariche che svolgono un ruolo principale dipendono dalla lunghezza d'onda della luce incidente). Si indichi con **s** lo spostamento provocato dal campo elettrico **E**: se N è il numero di elettroni (di carica e) per unità di volume, si ha allora

$$\mathbf{I} = \frac{1}{4\pi}\frac{\partial \mathbf{E}}{\partial t} + Ne\frac{\partial \mathbf{s}}{\partial t},$$

ammettendo che lo spostamento **s** soddisfi un'equazione differenziale del tipo

$$m\frac{\partial^2 \mathbf{s}}{\partial t^2} + r\frac{\partial \mathbf{s}}{\partial t} + \ell\mathbf{s} = \mathbf{E}e,$$

cioè che l'elettrone, per effetto del suo spostamento, sia sottoposto ad una forza di richiamo $-\ell\mathbf{s}$ e ad una forza di tipo attrito $-r(\partial\mathbf{s}/\partial t)$ (modello evidentemente molto schematico).

Per studiare le onde piane monocromatiche di frequenza ω si cerca un sistema di vettori **E**, **H**, **I**, **s**, che siano funzioni sinusoidali di frequenza ω. Dal momento che le equazioni sono lineari, si cerca una soluzione in $e^{i\omega t}$: la parte reale sarà la vera soluzione. Si ponga $\mathbf{s} = Ae^{i\omega t}$, dove A è indipendente da t, se ne deduce, quindi, dall'equazione differenziale soddisfatta da **s**, $(-m\omega^2 + ri\omega + \ell)\mathbf{s} = \mathbf{E}e$. Il corrispondente flusso elettrico **I** è

$$\mathbf{I} = \frac{\varepsilon}{4\pi}\frac{\partial \mathbf{E}}{\partial t}$$

dove ε è la costante complessa, che dipende dalla frequenza ω:

$$\varepsilon = 1 + \frac{4\pi N e^2}{\ell + ri\omega - m\omega^2}.$$

L'onda piana elettromagnetica di frequenza ω, soluzione dell'equazione di Maxwell, ha allora, come campo elettrico complesso, $\mathbf{E} = a\,e^{i\omega(t-kx)}$, $k = \omega\sqrt{\varepsilon/c}$. Si

definiscono il coefficiente di assorbimento K e l'indice reale ν per mezzo della radice quadrata del numero complesso ε, $\varepsilon = (\nu - iK)^2$; allora il campo elettrico reale è $\mathbf{E} = ae^{-(\omega K/c)x}\cos\omega[t - (\nu x/c)]$.

Generalmente si pone $\ell = m\omega_0^2$, e ω_0 si dice frequenza di assorbimento (è la frequenza di oscillazione libera degli elettroni): poiché il fattore r è sempre piccolo rispetto a $m\omega_0$, il coefficiente di assorbimento K, cioè la parte immaginaria del coefficiente dielettrico complesso ε, cessa di essere trascurabile solo per frequenze vicine alla frequenza di assorbimento ω_0.

L'onda illimitata $\mathbf{E} = ae^{i(\omega t - kx)}$ appena ottenuta corrisponde a un campo di oscillazioni nel punto $x = 0$, senza inizio né fine, $\mathbf{E}_0 = ae^{i\omega t}$.

Si consideri, ora, un segnale emesso a partire dal tempo $t = 0$, al punto $x = 0$, da un campo elettrico oscillante, di forma $f(t) = 0$ per $t < 0$, e $f(t) = \sin\alpha t$ per $t > 0$. Questo può essere considerato come una sovrapposizione di oscillazioni monocromatiche illimitate, che possono essere scritte nella forma

$$f(t) = \frac{1}{2\pi} Re \int \frac{e^{-i\omega t}}{\omega - \alpha}\, d\omega$$

(l'integrale deve essere preso lungo un cammino conveniente del piano complesso, per acquistare un significato non formale).

Dopo lo studio precedente, il campo elettrico corrispondente è, nel punto x e all'istante t,

$$\mathbf{E}(t, x) = \frac{1}{2\pi} Re \int \frac{e^{i(\omega t - kx)}}{\omega - \alpha}\, d\omega$$

dove $k = \omega\sqrt{\varepsilon/c}$.

Un argomento intuitivo, per ottenere la coincidenza delle velocità di segnale e di gruppo della frequenza α (tranne che questa sia nella «banda di assorbimento»), è che il contributo essenziale a $\mathbf{E}(t, x)$ provenga dai valori di ω vicini ad α, e che questo sia veramente importante solo se x e t sono tali che α verifichi la condizione segnalata nel paragrafo sulla «fase stazionaria», nel caso in cui k è reale, cioè lontano dalla frequenza di assorbimento.

Lo studio matematico conferma questo risultato e anche l'esistenza di precursori di debole ampiezza, che si propagano con velocità c.

Lo studio delle soluzioni discontinue delle equazioni di Maxwell in un mezzo dispersivo fornisce direttamente quest'ultimo risultato, che è valido anche vicino alle frequenze di assorbimento. Non si propaga alcun segnale con velocità superiore a quella della luce nel vuoto.

8. *Non-linearità.*

Nei paragrafi precedenti, si è ritenuta trascurabile l'azione dei fenomeni studiati su se stessi; li si è considerati come agenti in un mezzo nel quale la loro propagazione è additiva: questa ipotesi semplificatrice è valida solo per un'approssimazione piú o meno buona secondo la natura dei fenomeni considerati. In molti fenomeni reali, infatti, l'interazione del fenomeno con se stesso, a causa

della non-linearità delle leggi dinamiche, svolge un ruolo essenziale. Introduce, in particolare, un nuovo tipo di propagazione: quella degli urti.

Generalmente, s'intende con 'urto' la discontinuità di una grandezza fisica che appare, nella maggior parte dei casi, continua rispetto a una scala di misura convenzionale: in un problema di dinamica dei gas, un urto corrisponde a una discontinuità di velocità o di pressione. Nello studio eseguito sull'acustica (equazioni linearizzate) si è parlato di onde balistiche, o onde di Mach (prodotte da una sorgente a velocità supersonica), corrispondenti a un accumulo di rumore, che sono in effetti delle onde d'urto che si propagano con la velocità del suono, cioè con la velocità dei fronti d'onda usuali (che sono, in questo caso, discontinuità di accelerazione e di gradiente della pressione). Questi urti sono detti urti caratteristici, o urti deboli, in quanto sono di piccola ampiezza. Quando si tiene conto della non-linearità, si vede apparire un nuovo tipo di urto, la cui velocità di propagazione non è piú quella dei fronti d'onda.

La natura del fenomeno studiato è, spesso, profondamente modificata dopo l'urto e quindi le equazioni fin qui descritte non sono piú valide.

8.1. Problema del pistone.

Un modo intuitivo per prevedere la comparsa di onde d'urto è quello di osservare che, per equazioni non lineari, la velocità di propagazione dei fronti d'onda dipende dalla soluzione considerata. Per esempio, in un gas, essa cresce con l'aumentare della pressione: allora, una perturbazione si propagherà tanto piú velocemente quanto piú sarà grande, il che condurrà a una discontinuità.

Quest'ipotesi fu proposta già da Riemann, cui si deve la soluzione del problema del pistone, il quale illustra la nascita di un urto. Il suo metodo è ancora adesso utilizzato per risolvere problemi non lineari unidimensionali. Si dimostrerà rigorosamente, con un semplice esempio, la comparsa di urti, anche per piccoli valori, molto regolari, dei dati iniziali.

Si riprendano le equazioni unidimensionali della dinamica dei gas, supponendo costante per semplicità l'entropia:

$$\frac{\partial\rho}{\partial t}+\frac{\partial}{\partial x}(\rho u)=0$$

$$\rho\left(\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}\right)+\frac{\partial p}{\partial x}=0.$$

Si è visto che le equazioni linearizzate corrispondenti (per perturbazioni di piccola ampiezza) portano per p (e per u) a una stessa equazione del tipo di quella di d'Alembert, la cui soluzione generale è somma di un'onda progressiva $F(x-a_0t)$ e di un'onda regressiva $G(x+a_0t)$, dove a_0 è la velocità del suono nel mezzo non perturbato, $a_0^2=(dp/d\rho)_{\rho=\rho_0}$.

Si può estendere la definizione di onda progressiva, che si dice onda semplice, al caso non lineare, cosa che permette di risolvere il problema di un moto indotto da un pistone in un gas che riempia un tubo cilindrico. Si ricordi che

le equazioni dinamiche, tenuto conto dell'equazione di stato $p=p(\rho)$, implicano le due equazioni di evoluzione

$$\frac{\partial p}{\partial t}+(u+a)\frac{\partial p}{\partial x}+a\rho\left[\frac{\partial u}{\partial t}+(u+a)\frac{\partial u}{\partial x}\right]=0$$

$$\frac{\partial p}{\partial t}+(u-a)\frac{\partial p}{\partial x}-a\rho\left[\frac{\partial u}{\partial t}+(u-a)\frac{\partial u}{\partial x}\right]=0$$

in cui si è posto $a^2=dp/d\rho$.

La prima equazione di evoluzione dice che su una curva C_+ (che dipende dalla soluzione u, ρ) del piano degli (x, t) di equazione differenziale $dx/dt=u+a$, si ha

$$\frac{dp}{dt}+\rho a\frac{du}{dt}=0 \qquad \text{cioè } \frac{a\,d\rho}{\rho\,du}+1=0;$$

dunque, su una curva C_+,

$$\int\frac{a(\rho)\,d\rho}{\rho}+u=\text{costante} \qquad \text{su } C_+:\ \frac{dx}{dt}=u+a$$

La seconda equazione di evoluzione fornisce, con un analogo ragionamento,

$$\int\frac{a(\rho)\,d\rho}{\rho}-u=\text{costante} \qquad \text{su } C_-:\ \frac{dx}{dt}=u-a.$$

I primi membri sono detti «invarianti di Riemann».

Per confronto con la teoria linearizzata, in cui le curve C_+ e C_- sono sostituite dalle rette $x-a_0t=$ costante e $x+a_0t=$ costante, e in cui un'onda progressiva è funzione solo di $x-a_0t$, si dice onda semplice progressiva della teoria non lineare una soluzione delle equazioni dinamiche, il cui secondo invariante di Riemann è dovunque costante (cioè ha lo stesso valore sulle differenti curve C_-).

Si consideri un gas omogeneo in quiete all'istante iniziale $t=0$ che occupi lo spazio $x\geq 0$: la sua equazione di stato è $p=K\rho^\gamma$, dunque $a^2=K\gamma\rho^{\gamma-1}$. Gli invarianti di Riemann sono allora

$$\frac{2}{\gamma-1}a\pm u=\text{costante} \qquad \text{su } \frac{dx}{dt}=u\pm a.$$

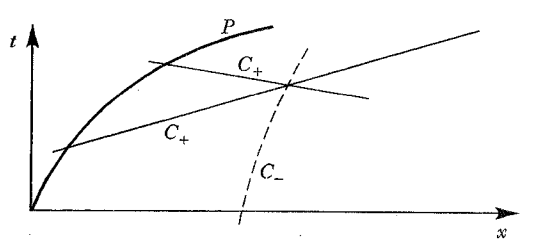


Figura 17.
Traiettoria di un pistone.

Un pistone posto nel punto di ascissa $x=0$, al tempo $t=0$ è spinto verso destra. La sua traiettoria nel piano degli (x, t) (fig. 17) $x=X(t)$ è una linea di flusso; si ha $dX/dt=u$ su $x=X(t)$. Si consideri la famiglia di caratteristiche C_-; si ha

$$\frac{2}{\gamma-1}a-u=\text{costante su ogni } C_-: \frac{dx}{dt}=u-a$$

all'istante $t=0$, e qualunque sia x, si è supposto $u=0$ e $\rho=\rho_0$, quindi $[2/(\gamma-1)]a-u=[2/(\gamma-1)]a_0$ è la stessa costante su tutte le curve C_- che incontrano l'asse ox, con $x\geq 0$. Le linee di flusso del fluido partono da questo semiasse per $t=0$ – lo stesso avviene per le curve C_- in quanto queste corrispondono a un piú piccolo valore di dx/dt – e si ha che il secondo invariante di Riemann è costante in tutta la regione D occupata dal gas nel corso del tempo.

Questo risultato, insieme alla considerazione del primo invariante di Riemann

$$\frac{2a}{\gamma-1}+u=\text{costante su ogni } C_+: \frac{dx}{dt}=u+a,$$

dimostra che a e u sono entrambe costanti su ogni curva C_+, cosa che si esprime nella forma

$$u=\text{costante su ogni } C_+: \frac{dx}{dt}=a_0+\frac{\gamma-1}{2}u.$$

Il valore di u su ciascuna di queste curve (che sono delle rette) è diverso poiché dipende dall'istante in cui incontra il pistone. Se lo incontra all'istante τ, il valore di u su questa curva C_+ è $u=\dot{X}(\tau)$. La sua equazione è, allora

$$x=X(\tau)+\left\{\left(a_0+\frac{\gamma-1}{2}\right)\dot{X}(\tau)\right\}(t-\tau).$$

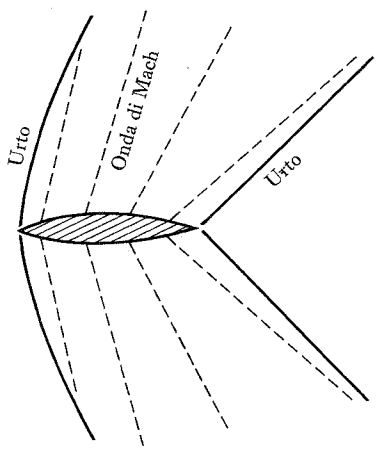


Figura 18.
Onde di Mach su un profilo alare in un flusso supersonico.

Al contrario, il valore di τ, in funzione di x e t, ricavato da questa equazione e riportato in $u=\dot{X}(\tau)$, fornisce il campo delle velocità $u(x, t)$ nel gas (e anche delle densità e pressioni).

Le curve C_+ sono rette di pendenza dx/dt crescente con u: si vede che se la velocità del pistone dX/dt cresce (caso della compressione) s'incontrano due curve che nascono da punti diversi e che forniscono due valori diversi delle incognite u e p in uno stesso punto: infatti, vicino a un tale punto, l'analisi fatta fin qui cessa di essere valida: non esistono soluzioni regolari e si genera un urto.

La figura 18 rappresenta le posizioni di alcune onde di Mach generate da differenti punti di un'ala di un aereo in volo supersonico, e di onde d'urto attaccate davanti e dietro agli spigoli. In altri casi, l'onda d'urto anteriore può essere staccata. Lo studio completo, che dipende dalla geometria dell'ala, è complesso ma importante per l'aeronautica: ha dato luogo a numerosi lavori teorici e sperimentali.

8.2. Solitoni.

Scott Russell scoprí sperimentalmente nel 1844 un fenomeno molto curioso, riguardante la propagazione delle onde sulla superficie dell'acqua. Mentre seguiva a cavallo il bordo di un canale, egli osservò un'onda isolata, di altezza apprezzabile, che si propagava senza deformazioni né smorzamenti, con velocità costante. Egli poté seguirla per piú di un chilometro. Non è possibile spiegare con la teoria linearizzata l'esistenza di un'onda simile, detta onda solitaria, in quanto contraria al concetto usuale di treni d'onda, originati da gruppi oscillanti che si smorzano nel tempo. È invece possibile spiegarla con una teoria non lineare; Boussinesq (1871) e Lord Rayleigh (1876) vi si sono applicati con l'ausilio di metodi approssimati. Nel 1895, Korteweg e Devries hanno dimostrato che le onde di gravità alla superficie di livelli d'acqua poco profondi, almeno in prima approssimazione, soddisfano l'equazione che ora si dice di Korteweg-Devries

$$\eta_t+\sigma\eta\eta_x+\eta_{xxx}=0 \qquad \eta_t=\frac{\partial\eta}{\partial t} \qquad \eta_x=\frac{\partial\eta}{\partial x}$$

dove σ è una costante e η indica il dislivello rispetto alla superficie in quiete. Si ottiene quest'equazione considerando onde progressive, cioè che hanno il primo invariante di Riemann costante, e considerando solo il primo termine non lineare (trascurando, cioè, il prodotto per η_x di termini di ordine $(\eta/h_0)^2$, dove h_0 è la profondità dell'acqua in quiete). Si possono ottenere molto semplicemente delle soluzioni esatte, che rappresentano delle onde solitarie, dette anche solitoni: con la sostituzione $\eta=(12/\sigma)(\log F)_{xx}$ si trasforma l'equazione di Korteweg-Devries nell'equazione quadratica omogenea

$$F(F_t+F_{xxx})_x-F_x(F_t+F_{xxx})+3(F_{xx}^2-F_xF_{xxx})=0.$$

Questa trasformazione è motivata dal fatto che l'equazione ottenuta ammette soluzioni esatte che sono esponenziali; $F=1+f$ con $f=\exp\{-\alpha(x-s)+\alpha^3t\}$ è una soluzione qualunque siano i parametri α e s; essa verifica $F_t+F_{xxx}=0$ e $F_{xx}^2-F_xF_{xxx}=0$.

Di conseguenza alle corrispondenti soluzioni esatte dell'equazione di Korteweg-Devries

$$\eta = \frac{12}{\sigma}\frac{\alpha^2 f}{(1+f)^2}$$

si dà il nome di solitoni: in un dato istante t, queste possiedono un massimo unico per $f=1$, cioè nel punto $x=s+\alpha^2 t$, e si smorzano rapidamente a partire da questo massimo. Quest'ultimo si sposta nel tempo con velocità costante α^2.

È anche possibile, con l'aiuto di soluzioni esatte dell'equazione di Korteweg-Devries, rappresentare l'interazione di due onde solitarie di velocità α_1^2 e α_2^2. Si cerchi una soluzione esatta sotto forma di serie $F=1+F^{(1)}+F^{(2)}+\ldots$ con $F^{(1)}=f_1+f_2$, $f_i=\exp[-\alpha_i(x-s_i)+\alpha_i^3 t]$, $i=1,2$. Si trova che la funzione

$$F=1+f_1+f_2+\frac{(\alpha_2-\alpha_1)^2}{(\alpha_2+\alpha_1)^2}f_1 f_2$$

è, anch'essa, una soluzione esatta dell'equazione trasformata di Korteweg-Devries. Questa soluzione può essere interpretata come descrittiva dell'interazione delle due onde solitarie. Si vede, infatti, scrivendo il dislivello η corrispondente a questa soluzione F, che corrisponde approssimativamente a un'onda solitaria di velocità α_1^2 quando f_1 è prossimo a 1, e f_2 è molto diversa. Piú precisamente, si trova che per $f_1 \simeq 1$ e $f_2 \ll 1$

$$\eta = \frac{12}{\sigma}\frac{\alpha_1^2 f_1^2}{(1+f_1)^2}$$

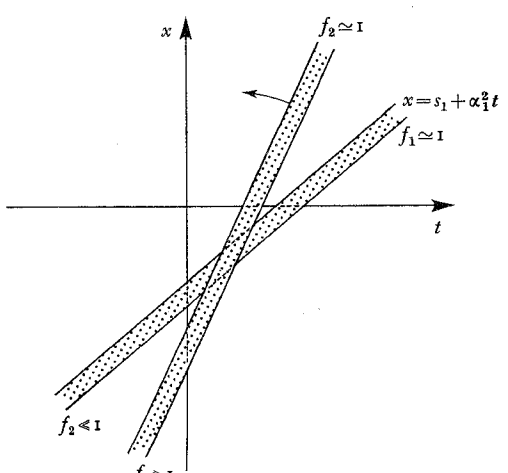


Figura 19.
Interazione di due onde solitarie.

e corrisponde quindi ad un'onda solitaria di velocità α_1^2. Per $f_1 \simeq 1$ e $f_2 \gg 1$, esso diventa

$$\eta = \frac{12}{\sigma}\frac{\alpha_1^2 \tilde{f}_1}{(1+\tilde{f}_1)^2}$$

dove

$$\tilde{f}_1 = \left(\frac{\alpha_2-\alpha_1}{\alpha_2+\alpha_1}\right)^2 f_1$$

e corrisponde ancora ad un'onda solitaria di velocità α_1^2 il cui profilo, però, è spostato rispetto alla precedente, nella direzione ox, della quantità

$$-\frac{2}{\alpha_1}\log\left(\frac{\alpha_2+\alpha_1}{\alpha_2-\alpha_1}\right)^2.$$

Un simile risultato è valido anche quando f_2 è prossimo a 1 e f_1 è grande o piccolo. La regione in cui sia f_2 sia f_1 sono vicini a 1, è la regione d'interazione. Si supponga, per fissare le idee, $\alpha_2 > \alpha_1$ (fig. 19). In un lontano passato ($t \to -\infty$), il dislivello corrispondente a $F=1+f_1+f_2+[(\alpha_2-\alpha_1)/(\alpha_2+\alpha_1)]^2 f_1 f_2$, descrive due onde solitarie, una di velocità α_1^2, situata in una regione in cui $f_1 \simeq 1$ e $f_2 \ll 1$, con dislivello, dunque, equivalente a η_1, situato nelle vicinanze di $x=s_1+\alpha_1^2 t$; l'altra di velocità α_2^2, situata in una regione in cui $f_2 \simeq 1$ e $f_1 \gg 1$, con dislivello equivalente a η_2 situato nelle vicinanze di $x=s_2-(1/\alpha_2)\log[(\alpha_2+\alpha_1)/(\alpha_2-\alpha_1)]^2+$ $+\alpha_2^2 t$. Quando t diventa molto grande e positivo, un analogo studio dimostra che le due onde solitarie emergono dalla loro interazione con la stessa forma e velocità, l'onda α_2 è ormai in testa; la sola traccia della collisione è uno spostamento in avanti di $(2/\alpha_2^2)\log[(\alpha_2+\alpha_1)/(\alpha_2-\alpha_1)]$ per l'onda α_2, e uno spostamento regressivo: $(1/\alpha_1)\log[(\alpha_2+\alpha_1)/(\alpha_2-\alpha_1)]$ per l'onda α_1.

Si trova cosí il notevole risultato che se numerose onde solitarie, convenientemente spaziate, s'incrociano, queste possono uscire dalla loro interazione ritrovando la loro forma e la loro velocità originali. Il solo effetto della loro interazione è uno spostamento costante rispetto alle posizioni che queste avrebbero avuto senza interazione. L'analogia con la collisione delle particelle è sorprendente e ha risvegliato un grande interesse collegato con le teorie quantistiche in cui il doppio aspetto corpuscolare e ondulatorio è sempre presente.

La ricerca di solitoni, soluzioni particolari notevoli di equazioni alle derivate parziali non lineari, è ancora un campo molto attivo della fisica matematica, in particolare dopo l'introduzione del metodo della diffusione inversa (*inverse scattering*). [Y. CH.-B.].

I fenomeni (cfr. **fenomeno**) oscillatori sono numerosi: dalle onde marine a quelle sonore alla propagazione della **luce**, il **moto** si trasmette da un punto a un punto vicino (cfr. anche **energia, interazione**) senza trasporto di **materia** (per questo cfr. anche **particella, quanti**). Lo studio di questi fenomeni, dalle piú semplici vibrazioni di una corda a manifestazioni piú complesse, è condotto coi metodi delle equazioni differenziali (cfr. **differenziale**, e anche **infinitesimale, funzioni, stabilità/instabilità**) e rappresenta un importante settore della fisica matematica (cfr. **fisica, spazio-tempo**).

na), so-
co-

Quanti

1. *Discreto e continuo.*

La carica semantica ormai connessa con il lemma 'quanti' investe un ampio spettro di situazioni e di contesti disciplinari: in questi importa qui cogliere un paradigma comune. Persino nell'uso meno tecnico possibile della parola, quello connesso con la domanda «quanti?» intesa ad ottenere una risposta numerica, si riconosce nella sua elementarità un tale paradigma: la domanda individua immediatamente una demarcazione tra risposte numeriche possibili e risposte non possibili; chiunque non sia tradito da troppo impegnative conoscenze matematiche è pronto ad accettare – tra le risposte possibili – numeri interi, eventualmente negativi, e forse anche numeri razionali. Si possono volere due fette di torta, si può asserire di possedere -3 milioni di lire, intendendo con questo l'esistenza di un tale debito, si può forse anche desiderare di possedere $8/13$ della stessa torta, pur di essere forniti di un'opportuna capacità di ripartirla in parti uguali. Sembra invece molto arduo, a parte la bizzarria della cosa, desiderare $\pi/13$ di torta, anzitutto perché non è possibile fornire – pur provvisti dell'ovvia conoscenza che π è il rapporto costante tra una circonferenza e il suo diametro – alcuna procedura finita in grado di produrre una tale «frazione».

Sono dunque ipotizzabili due classi di risposte – alla domanda «quanti?» – a prima vista nettamente distinguibili e con un livello di significanza tra loro molto diverso. Già in un contesto cosí semplificato si riconoscono alcuni elementi di quella polarità discreto/continuo che ha svolto e svolge nella storia del pensiero un ruolo cosí rilevante.

Sulle complicate e alterne vicende, sia sul versante matematico sia su quello delle scienze della natura, attraversate da tale contraddizione molto è stato scritto e ancora molto si dibatte.

Non sarà quindi inutile spendere qualche parola su alcuni punti significativi della storia della scienza nel mondo occidentale. Va anzitutto detto che, malgrado la continua interazione tra i due terreni, converrà tenere distinto l'ambito della matematica da quello delle scienze naturali. Se parlare di matematica del discreto significa alludere all'aritmetica dei numeri naturali, sembra chiaro che ad essa spetta la priorità temporale nell'elaborazione della matematica e che l'esigenza di «riempire gli spazi vuoti» tra i numeri naturali (o eventualmente tra i numeri razionali) – in sé sufficienti per *contare* – sia venuta piuttosto dal bisogno di *misurare*; la diagonale del quadrato non si misura con un numero espresso sotto forma di frazione in termini del lato. Appare legittimo assumere che l'esigenza di arricchire la matematica con un insieme di numeri piú completo ed efficiente che non l'insieme dei naturali o dei razionali, sia venuta da un'intuizione fisica assai primitiva: quella che l'estensione dei corpi non presenta, al livello dell'osservazione macroscopica, tranne che ai suoi bordi, alcuna «soluzione di continuità».

Non va d'altra parte dimenticato che è presente anche nella fisica, e fin dall'antichità, un filone di pensiero che si appoggia – per le proprie spiegazioni scientifiche dei fenomeni naturali – ad una struttura discreta della materia a livello microscopico: l'atomismo di Leucippo e di Democrito, e di Epicuro e di Lucrezio, ebbe significati e connotazioni che andarono certamente oltre il periodo limitato del suo primo apparire, e rappresentò un'alternativa alle concezioni continuiste fino al secolo scorso; un'alternativa, come si sa, in ultima istanza vincente. Certo è tuttavia che – nella misura in cui le osservazioni sperimentali rimasero fino a tutto il secolo scorso di tipo macroscopico – l'atomismo rimase una ipotesi esplicativa, in apparenza precisa e articolata, in realtà vaga e adattabile a vari modi di argomentare e quindi difficilmente suscettibile di limpide corroborazioni o falsificazioni. Perfino in periodo medievale, osserva Dijksterhuis [1950] parlando dell'opera enciclopedica di Isidoro e Rabano Mauro, «l'antico atomismo era in grado di fornire, a coloro che non facevano troppe domande ma si accontentavano di spostare le difficoltà dal macrocosmo al microcosmo, spiegazioni apparentemente molto plausibili di moltissimi fenomeni fisici comuni. Questo fatto fu sfruttato, ad esempio, da San Gerolamo nella sua esegesi della Scrittura» (trad. it. p. 139). D'altra parte, delle varie determinazioni della contraddizione discreto/continuo, questa dell'atomismo – dibattuta con fasi incerte fino a quasi tutto l'Ottocento (cfr. ad esempio l'articolo «Forza/campo» in questa stessa *Enciclopedia*) – non può certo dirsi coincidere con quella che verrà segnata profondamente dalla rottura quantistica del 1900: rottura che invece si riferisce a un atteggiamento continuista di piú ampio ed «astratto» respiro. E cioè quello che aveva sempre piú o meno tacitamente suggerito che (una volta che la matematica – attraverso un processo assai lungo e non certo facilmente schematizzabile – aveva riempito i «buchi» di cui si diceva e aveva quindi esibito una struttura assiomatica rigorosa per il campo dei numeri reali) le variabili matematiche che rappresentano grandezze fisiche *prima facie* continue, potessero per l'appunto assumere tutti i valori reali compresi in un determinato ***range***: cosí per la posizione, per il momento lineare, per l'energia, ecc.

Questa convinzione che in natura tutte le possibilità di valori per certe variabili sono in realtà attuate o attuabili, strettamente connessa con quell'altra certezza che «natura non facit saltus» ha in verità radici assai antiche e praticamente unanimi. Si ricorderà l'opinione di Aristotele, secondo cui v'è continuità persino tra la sfera del vivente e quella del non-vivente: «La natura passa per gradi cosí impercettibili dagli esseri inanimati agli animali, che tale continuità rende impossibile scorgere il confine fra i due campi e decidere a quale di essi appartenga la zona intermedia» [*Ricerche sugli animali*, 588b]. Non è certo possibile citare qui le innumeri testimonianze del progressivo, e automatico, adattamento della fisica – e dunque dei *range* di valori assumibili dalle grandezze fisiche – agli sviluppi e ai bisogni della matematica, in particolare alla costruzione e all'assiomatizzazione dei numeri reali. Valga per molte l'argomentazione a favore di una struttura continua dello spazio-tempo che si trova in Hegel [1830]: egli vede in tale continuità l'unica possibilità di risolvere i paradossi zenoniani – «Il movimento entra nell'antinomia zenoniana, che è insolubile se i luoghi vengono *isolati* come punti spaziali, e i momenti del tempo come punti temporali; e la soluzione dell'antinomia, cioè il movimento, è da comprendere solo cosí, che spazio e tempo in sé sono continui» (trad. it. pp. 282-83) –, non avvedendosi in verità che in uno spazio-tempo discreto, cioè fatto di una quantità localmente finita di punti, il ragionamento che impedisce ad Achille di raggiungere la tartaruga perde validità dopo un numero finito di passi, e dunque il paradosso scompare comunque. La moderna consapevolezza dell'esistenza di strutture matematiche discrete (fondate ad esempio sul concetto di campo finito) mostra come non vi sia nulla di assai cogente nella scelta del continuo piuttosto che in quella del discreto, pur naturalmente di considerare, ad esempio per descrivere lo spazio-tempo, strutture discrete con un numero, sí finito, ma sufficientemente alto di elementi. Quale sia del resto lo strumento matematico migliore per una teoria che descrive un determinato ambito di fenomeni è materia di un giudizio necessariamente a posteriori, rispetto alla teoria e alle sue verifiche, e comunque sempre problematico. Del resto «rimane pur certo che il continuo della nostra intuizione e il continuo matematico non coincidono; tra i due si apre un profondo abisso. Ma ciononostante siamo spinti da considerazioni ragionevoli quando, nel nostro desiderio di comprendere il mondo, passiamo dall'uno all'altro» afferma Weyl [1917], e aggiunge ambiguamente in nota: «Cosí, per esempio, non potremmo associare la continuità allo schema dei numeri interi: ciò non dipende dal nostro arbitrio. Tuttavia, chi può sapere quante cose nel campo della fisica riposano ancora nel seno del futuro e della teoria dei quanti!» (trad. it. p. 145). Stante tuttavia l'unanime e indiscusso affidamento continuista che caratterizza – nel senso sopra precisato – la storia della scienza fino a tutto il secolo scorso, le nuove proposte con cui si annunzia la fisica del nostro secolo appaiono certamente una rottura di enormi proporzioni. Può ben dirsi la rottura di un consolidato paradigma (si veda l'omonimo articolo in questa stessa *Enciclopedia*), anche se Kuhn stesso, nel suo piú recente e accurato studio dedicato a questo «caso storico» [1978], sembra non servirsi quasi piú di questa terminologia. Né si deve pensare che una tale rottura si sia rapidamente affermata ed abbia immediatamente segnato il pensiero successivo, se ancora nel 1918, quando ormai, come si vedrà, una prima fase della storia della «old quantum theory» si era conclusa, Wittgenstein [1922] poteva scrivere: «Tutte le proposizioni come il principio di ragion sufficiente, di continuità nella natura, del minimo sforzo nella natura, etc. etc., sono tutte intuizioni *a priori* sulla possibile formulazione delle proposizioni della scienza» (trad. it. p. 75). Quanto poco si trattasse di un'intuizione a priori nel senso wittgensteiniano del termine [cfr. *ibid.*, §§ 6.32 e 6.3211] era cominciato ad apparire fin dall'estate del 1900.

2. *Le nuvole nere e il ruolo di Planck.*

Come ricorda Broglie nella sua tesi di dottorato [1924, p. 13], alla fine del XIX secolo Lord Kelvin «annunziava che due nuvole nere apparivano minacciose all'orizzonte della fisica» [alludendo a Kelvin 1884, p. 491]. La prima nuvola

era rappresentata dal complesso delle esperienze del solo Michelson e di Michelson e Morley poi che, a dispetto delle critiche e delle reinterpretazioni di Lorentz, sembravano introdurre sempre nuove contraddizioni nella teoria della propagazione delle onde elettromagnetiche attraverso l'etere. La seconda minaccia era costituita dalla cosiddetta catastrofe ultravioletta che caratterizzava la teoria esistente a proposito dell'emissione di radiazione da parte di un corpo nero. Osserva poco oltre Broglie: «Nei primi anni del XX secolo le due nuvole di Lord Kelvin si sono, se cosí posso dire, condensate l'una nella teoria della relatività e l'altra nella teoria dei quanti» [1924, p. 14].

Fin dalle sue origini, la teoria quantistica fu fonte di dibattito e disaccordo tra gli storici e i filosofi della scienza: non v'è neppure unanimità sulla sua data di nascita, come si vedrà; la prima cosa che occorre mettere in luce qui è la natura della seconda nuvola nera che si addensava negli ultimi anni del secolo sul capo dei fisici europei. Con gli apporti di Maxwell e di Boltzmann nel campo oggi noto come meccanica statistica cominciò ad emergere l'importanza di una teoria che fosse in grado di applicare la termodinamica classica ai fenomeni connessi con l'emissione e l'assorbimento di radiazione elettromagnetica. Se si tiene presente che tale termine abbraccia l'intero spettro di radiazioni, con frequenza ν che può assumere qualsiasi valore reale positivo, dalle onde radio alle onde luminose (visibili), alle onde termiche (spettro infrarosso), ai raggi X e γ, ecc., e se si tiene conto che ad ogni radiazione è connessa una certa quantità di energia, acquisteranno un senso assai generale la definizione di potere emissivo e di potere assorbente di un corpo qualsiasi; con il primo termine s'intende l'energia irradiata dal corpo in esame nell'unità di tempo e con potere assorbente la frazione dell'energia radiante incidente sul corpo che viene da questi assorbita. Risulta fondamentale la considerazione di un corpo ideale – detto corpo nero – per il quale il potere assorbente è uguale a 1. In base ad una legge fondata soltanto su considerazioni termodinamiche e stabilita da Kirchhoff fin dal 1859, il rapporto tra il potere emissivo e il potere assorbente di un corpo dipende soltanto dalla sua temperatura e non dalla sua particolare natura; si conclude che per il corpo nero, il cui potere assorbente è 1, il potere emissivo dipende solo dalla temperatura. La funzione $u(\nu,T)$ che esprime l'energia irradiata per unità di tempo, al variare della temperatura del corpo nero e della frequenza della radiazione emessa, fu oggetto negli ultimi decenni del secolo scorso di ricerche assai approfondite; le principali tappe che qui interessa ricordare sono le seguenti. Anzitutto va menzionata la legge detta di Stefan-Boltzmann, scoperta nel 1879 da Stefan e in seguito dimostrata da Boltzmann sfruttando il concetto di pressione di radiazione basato sulla teoria di Maxwell, secondo la quale l'energia totale emessa dal corpo nero è proporzionale alla quarta potenza della temperatura (si ricordi che sempre ci si riferisce, qui e nel seguito, alla temperatura assoluta): $u(T)=\sigma T^4$, ove $u(T)=\int_0^\infty u(\nu,T)\,d\nu$ è l'energia complessivamente emessa alla temperatura T, tenendo conto di tutte le frequenze possibili.

Wien trovò in seguito (1896) la cosiddetta legge di spostamento da cui segue la stessa legge di Stefan-Boltzmann: questa può essere enunciata dicendo che la funzione u di cui si parla è della forma $u(\nu,T)=\nu^3F(\nu/T)$, ove F è una funzione

che non può essere determinata sulla base di sole considerazioni termodinamiche. La termodinamica impone soltanto, in altre parole, che u/ν^3 dipenda dal rapporto ν/T. È opportuno sottolineare fin d'ora che il fatto che tali leggi fossero ricavate sfruttando soltando i principî base della termodinamica, ritenuti assolutamente certi, fece sí che esse non fossero in nessun caso poi messe in discussione; e ciò a differenza di altri dogmi della fisica classica, in apparenza altrettanto fondamentali, che invece, come si dirà, furono piú volte modificati o rovesciati. Se si scrive l'energia irradiata in funzione della lunghezza d'onda, $\lambda=c/\nu$ (c è la velocità della luce nel mezzo) invece che della frequenza, si ha $u_\lambda(\lambda,T)=(c^4/\lambda^5)F(c/\lambda T)$; se ora ci si chiede per quali valori di λ tale funzione ha un estremo (in questo caso un massimo) – il che significa chiedersi, per una fissata temperatura T, qual è la lunghezza d'onda piú favorita nell'emissione – si trova, annullando la derivata $\partial u/\partial\lambda|_{T=\text{cost}}$, un'espressione del tipo: $(c/\lambda T)F'(c/\lambda T)+5F(c/\lambda T)=0$; cioè, qualunque sia la forma della F, un'equazione nel prodotto λT; ciò significa in particolare che, se si pensa ora di far variare la temperatura T, il punto λ_m in cui la u_λ assume il suo massimo si sposta (di qui il nome «legge di spostamento») verso valori maggiori se T diminuisce e minori se T aumenta; si può equivalentemente affermare che la frequenza cui si ha il massimo è direttamente proporzionale alla temperatura considerata.

Considerazioni basate solo sulla termodinamica classica non possono d'altra parte portare alla determinazione precisa della funzione F, per conoscere la quale occorre ipotizzare un meccanismo particolare per l'emissione: un modello che tipicamente venne proposto come il piú naturale emettitore di radiazione elettromagnetica fu quello costituito da un oscillatore armonico carico, di frequenza ν; detta $\bar{E}$ l'energia di un tale oscillatore mediata su tempi grandi rispetto al periodo di oscillazione, si trova per l'energia emessa $u_\nu=(8\pi\nu^2/c^3)\bar{E}$; tutta la questione si riduce allora a calcolare $\bar{E}$ in funzione della temperatura. Qui è il punto cruciale del calcolo: per calcolare $\bar{E}$, due ipotesi vanno sfruttate: *a*) che tutti i valori dell'energia siano a priori possibili; *b*) che la probabilità di ciascuno di essi, indicato con E, in accordo con un fondamentale teorema di meccanica statistica dovuto a Boltzmann, sia proporzionale a $p(E)=e^{-E/kT}$, essendo k la costante di Boltzmann (pari a circa $1{,}38\cdot10^{-16}$ erg deg^{-1}); il calcolo effettuato sulla base di queste due ipotesi è assai semplice e fornisce (scrivendo per brevità $\beta=1/kT$)

$$\bar{E}=\frac{\int_0^\infty Ee^{-\beta E}\,dE}{\int_0^\infty E^{-\beta E}\,dE}=-\frac{d}{d\beta}\ln\int_0^\infty e^{-\beta E}\,dE=-\frac{d}{d\beta}\ln\frac{1}{\beta}=\beta^{-1}=kT;$$

col che l'espressione per u_ν diventa $u_\nu=(8\pi\nu^2/c^3)kT$; questa legge, nota sotto il nome di Rayleigh e Jeans (ma dovuta essenzialmente al primo), è ovviamente in accordo con la prescrizione di Wien, pur di assumere $F(\nu/T)=(8\pi k/c^3)(\nu/T)^{-1}$ ed è, per piccoli valori della frequenza ν, in ottimo accordo con i dati sperimentali. Si riconobbe tuttavia immediatamente che essa non poteva valere per ogni frequenza ν per l'ottima ragione che $\int_0^\infty u_\nu\,d\nu$ non esiste finito: si ha cioè che l'emissione di radiazione, per una fissata temperatura T, dovrebbe aumentare

indefinitamente con la frequenza, e ciò in palese contraddizione con l'esperienza, oltre che con il principio di conservazione dell'energia. Rayleigh stesso propose, in conclusione alla sua memoria in proposito [1900], di introdurre il fattore esponenziale $u(\nu,T)\simeq\nu^2 T e^{-\beta\nu/T}$, senza tuttavia alcuna giustificazione teorica. Si trattava di uno dei molti tentativi (cinque ne vengono citati da Rubens e Kurlbaum [1900]) di mescolare il ragionamento fisico con un *guessing* matematico per trovare formule confermate dagli esperimenti.

È questa dunque la nuvola nera di cui si parlava; essa divenne nota sotto il nome di «catastrofe ultravioletta», poiché, rispetto allo spettro visibile, la regione ultravioletta è quella appunto delle alte frequenze.

La soluzione che venne infine offerta da Planck fu la seguente. Egli osservò che se l'ipotesi *a*) viene sostituita con l'ipotesi *a'*) – i valori permessi per l'energia degli oscillatori non sono tutti i numeri reali positivi ma solo multipli interi di un valore minimo E_0, piccolo ma finito –, mantenendo l'ipotesi *b*), il calcolo riportato sopra si modifica profondamente: agl'integrali si sostituiscono le somme, o per meglio dire, le serie; si ha infatti, per il calcolo di $\bar{E}$,

$$\bar{E}=\frac{\sum_{n=0}^{\infty} nE_0 e^{-\beta nE_0}}{\sum_{n=0}^{\infty} e^{-\beta nE_0}}=-\frac{d}{d\beta}\ln\sum_{n=0}^{\infty} e^{-\beta nE_0}=-\frac{d}{d\beta}\ln\frac{1}{1-e^{-\beta E_0}}=\frac{E_0}{e^{\beta E_0}-1},$$

cosí che la formula che fornisce u_ν diventa

$$u_\nu(\nu,T)=\frac{8\pi\nu^2}{c^3}\frac{E_0}{e^{\beta E_0}-1}=\frac{8\pi\nu^2}{c^3}\frac{E_0}{e^{E_0/kT}-1};$$

va ora osservato che la legge di Wien non è qui automaticamente soddisfatta; occorre fare un'ipotesi sul «quanto» E_0, infatti l'unico modo perché u_ν/ν^3 si possa scrivere come funzione soltanto del rapporto ν/T è che sia $E_0=h\nu$, con una certa costante di proporzionalità h; costante che non deve dipendere dal particolare modello adottato e che deve quindi possedere un significato universale; la legge di Planck per l'emissione di radiazione elettromagnetica da parte di un corpo nero assume quindi la forma

$$u_\nu(\nu,T)=\frac{8\pi\nu^3 h}{c^3}\frac{1}{e^{h\nu/kT}-1};$$

si noti in particolare che per piccole frequenze ($h\nu/kT\ll 1$) una buona approssimazione di tale formula è data da

$$u'_\nu(\nu,T)=\frac{8\pi\nu^3 h}{c^3}\frac{kT}{h\nu}=\frac{8\pi\nu^2}{c^3}kT,$$

che corrisponde alla formula ricordata di Rayleigh e Jeans; viceversa, per alti valori di $h\nu/kT$, una buona approssimazione è

$$u''_\nu(\nu,T)=\frac{8\pi\nu^3 h}{c^3}e^{-h\nu/kT},$$

cioè la formula scritta da Wien nel tentativo di spiegare il comportamento dell'emissione ad alta frequenza. La legge trovata da Planck possiede dunque alcune proprietà cruciali: coincide, nei loro *range* di applicazione, con leggi trovate in precedenza; fuori da questi dà una previsione in pieno accordo con l'esperienza. È della forma voluta dalla legge di spostamento di Wien e dà luogo a una u_ν integrabile nell'intervallo $[0,\infty)$, rispettando quindi anche la legge di Stefan-Boltzmann; se infatti si integra u_ν su $[0,\infty)$, si trova $\sigma=(2\pi^5k^4)/(15c^2h^3)\simeq$ $\simeq 5{,}67\cdot 10^{-5}$ erg cm^{-2} sec^{-1} deg^{-4}, valore in ottimo accordo con l'esperienza.

L'aver mostrato questa derivazione con qualche particolare dovrebbe permettere di comprendere in cosa consiste e quanto sia cruciale la differenza tra l'ipotesi *a*) e l'ipotesi *a'*); nessuna delle due è, a stretto rigore, piú motivata dell'altra, anche se la prima ha alle spalle, come s'è detto, secoli di autorevole tradizione. Resta da capire cosa spinse Planck a imboccare la strada *a'*) e con quanta coscienza innovativa egli lo fece. Certo è che egli nel corso del 1900 si dedicò a questo problema con grande intensità: dapprima [Planck 1900*a*, *b*], basandosi sull'ipotesi classica di connessione tra entropia e probabilità, derivò di fatto la legge di Wien; quando diventò chiaro che tale legge per basse frequenze era fortemente inadeguata – decisivi furono a questo proposito i risultati sperimentali di Kurlbaum, Rubens, Lummer, Pringsheim e Paschen –, Planck provò a modificare, in maniera semplice, la relazione detta. Cosí egli si esprime, dopo aver constatato che la legge $d^2S/dE^2=\text{cost}/E$ (essendo S l'entropia di un oscillatore che interagisce con la radiazione elettromagnetica, ed E la sua energia vibrazionale) porta appunto alla legge di Wien, non adatta alle basse frequenze: «Tuttavia, io potei considerare la possibilità, benché non facilmente comprensibile e in ogni caso difficile da dimostrare, che l'espressione che compare a sinistra nell'equazione possa non avere il significato generale che io le attribuivo in un primo tempo... Seguendo questo ragionamento ho infine cominciato a costruire espressioni completamente arbitrarie per l'entropia che, anche se piú complicate dell'espressione di Wien, sembrano tuttavia soddisfare altrettanto completamente tutte le richieste della termodinamica e della teoria elettromagnetica. Fui specialmente attirato da una delle espressioni cosí costruite, che è quasi altrettanto semplice dell'espressione di Wien e che merita di venir approfondita perché l'espressione di Wien non è sufficiente per un accordo con tutte le osservazioni. Otteniamo tale espressione ponendo $d^2S/dE^2=\alpha/E(\beta+E)$. È di gran lunga la piú semplice di tutte le espressioni che portano ad avere S come funzione logaritmica di E...» [1900*c*, pp. 203-4] e porta per l'appunto all'espressione vista piú sopra. Val la pena notare a questo punto che: 1) questa è l'occasione in cui per la prima volta compare la «giusta» legge che fornisce l'energia emessa in funzione della temperatura e della frequenza. Questa sua prima derivazione è del tutto ad hoc, priva cioè di ogni sostegno teorico indipendente; Planck stesso, come si è visto, è molto esplicito su questo punto: egli prova varie espressioni finché ne trova una che funzioni; 2) un criterio che lo guida è quello della semplicità, piú volte sottolineato; 3) nessuna menzione è fatta della possibilità di considerare solo quantità di energia in qualche modo discrete. Va forse aggiunto che in una ricostruzione assai posteriore della vicenda [1943, p. 156] Planck ad-

duce qualche motivazione per l'introduzione del termine a secondo membro della formula ora vista; ciò non sembra comunque sufficiente a infirmare le considerazioni esposte.

Il problema di *giustificare* l'espressione indovinata per u_ν assorbí Planck nei due mesi compresi tra il 19 ottobre, giorno in cui comunicò, in una seduta della Società tedesca di fisica di Berlino, la sua memoria [1900*c*], e il 14 dicembre, giorno in cui, nella stessa sede, presentò un brillante lavoro (cui usualmente vien fatto riferimento come punto d'inizio della teoria dei quanti) nel quale riusciva ad impiantare una costruzione teorica adatta a giustificare la nuova legge.

Planck era interessato, come s'è visto, a calcolare l'entropia del sistema di oscillatori e ciò implicava calcolare la probabilità di certe configurazioni energetiche rispetto ad altre; ovvero era interessato a calcolare in quanti modi possibili l'energia totale si poteva distribuire tra i vari oscillatori. Eseguire un tale calcolo tenendo valori continui per l'energia sembrava impossibile. Per capire di che si tratta, si supponga di voler calcolare la probabilità che, disegnando a caso un raggio in una certa circonferenza, questo cada entro un prefissato settore, si pensi per semplicità di 90°; se si tiene presente che la probabilità è data, nella sua accezione piú elementare, dal rapporto fra il numero dei casi favorevoli e il numero dei casi possibili, si avranno forti difficoltà ad effettuare il calcolo, visto che entrambi i numeri (casi favorevoli e casi possibili) sono infiniti; per poter procedere (ed ottenere il valore intuitivamente ovvio di 1/4) si può convenire che non tutti i raggi sono possibili, ma solo quelli che cadono in un numero molto grande ma finito di tacche poste ad intervalli regolari lungo la circonferenza; in tal modo si potrà ripristinare il metodo di calcolo detto ed ottenere effettivamente 1/4, e s'immaginerà anche che, se le tacche segnate sono abbastanza fitte, non si sia commesso, adottando questo trucco, alcun apprezzabile errore. Nel limite in cui il numero di tacche tende ad infinito, la probabilità rimane 1/4 e non vi è alcun problema ad assumere questo valore come risposta alla domanda iniziale. Il trucco venne adottato da Planck per calcolare la dipendenza dell'entropia del sistema dall'energia vibrazionale. « Dobbiamo ora dare la distribuzione dell'energia sui risuonatori separati di ogni gruppo, e prima di tutto la distribuzione dell'energia E sugli N risuonatori di frequenza ν. Se E è considerata una quantità infinitamente divisibile, la distribuzione è possibile in infiniti modi. Noi tuttavia consideriamo – ed è questo il punto piú essenziale di tutto il calcolo – che E sia composta di un numero ben definito di parti uguali e usiamo a tale scopo la costante di natura $h = 6{,}55 \cdot 10^{-27}$ erg sec... È chiaro che la distribuzione di P elementi di energia su N risuonatori può aver luogo soltanto in un numero di modi finito e ben precisato » [1900*d*, pp. 239-40].

Questa schematica descrizione già mostra come il primo apparire dell'ipotesi dei quanti non fosse basato sull'emergere di nuovi grandi orizzonti concettuali, ma su piú modesti terreni connessi con l'operare concreto del ricercatore, brillante certo, ma sempre legato alla soluzione di un particolare problema. Sembra del resto naturale che Planck non abbia attribuito, al suo primo apparire, all'ipotesi che l'energia possa venire scambiata soltanto in quantità discrete, quel valore rivoluzionario che doveva a poco a poco assumere, visto anche il ruolo strumentale ricoperto da tale ipotesi nei lavori originali del fisico berlinese. Kuhn [1978] sostiene, con grande abbondanza di documentazione storica, la tesi che la rottura abbia preso corpo davvero soltanto qualche anno dopo il 1900, e cioè essenzialmente grazie all'interpretazione di Ehrenfest ed Einstein, e al successivo suggello di Lorentz, dell'ipotesi di Planck; certo questa tesi non trova unanimi consensi. Klein [1979], ad esempio, esprime un sostanziale e motivato disaccordo. Non v'è dubbio comunque che « la storia di quest'avvenimento è tanto piú notevole in quanto, almeno nella sua prima fase, non fu assolutamente influenzato da scoperte clamorose di natura sperimentale » [Einstein 1946, trad. it. p. 21].

Sembra in ogni modo sicuro che la coscienza di una rottura con la teoria classica – la parte meglio stabilita della fisica del tempo – non sorse repentinamente nel Planck del 14 dicembre 1900; tale coscienza emerse con molta fatica tra i fisici della nuova generazione, di mano in mano che risultati teorici e sperimentali, relativi all'applicazione della nuova ipotesi ad altri ambiti che non fossero quelli della teoria del corpo nero, acquistava sempre piú rilevanza.

3. *La rottura si radicalizza: Bohr e Einstein.*

L'ipotesi di una struttura discreta per l'energia di alcuni sistemi fisici, che cominciò ad affermarsi decisamente verso la fine del primo decennio del secolo, fu, come si è accennato, un po' alla volta applicata a differenti settori della ricerca fisica. Un argomento chiave, che costituí poi uno dei pilastri dell'intera teoria quantistica, fu quello della struttura dell'atomo. Un notevole numero di esperienze attendeva di venire inquadrato in una teoria coerente della costituzione atomica. Rutherford aveva ipotizzato che l'atomo consistesse di un nucleo centrale dotato di una carica pari a $+Ne$ o a $-Ne$ (ove N è un intero positivo ed e la carica dell'elettrone) e da una sfera circostante uniformemente dotata di carica opposta ($-Ne$ oppure $+Ne$ rispettivamente), essendo il raggio di tale sfera dell'ordine di 10^{-8} cm. Il modello di Rutherford permise di spiegare numerose esperienze. Si trattava però di andare oltre l'aspetto meramente modellistico e di trovare qualche nuova legge fondamentale che spiegasse in maniera complessiva il comportamento degli atomi e in particolare della loro interazione con la radiazione elettromagnetica; in altre parole i processi di emissione e di assorbimento di radiazione. Fu su questo punto che l'opera del danese Niels Bohr si rivelò a poco a poco con caratteristiche profondamente innovatrici. « La sintesi del modello atomico di Rutherford con l'ipotesi dei quanti di Planck fu la grande conquista di Niels Bohr », scrive Waerden [1967, p. 2], e in quale maniera fu proprio cosí; in una serie di lavori pubblicati tra gli anni 1913 e 1915 Bohr aveva un po' alla volta ottenuto un'interpretazione ragionevolmente semplice delle leggi che governano gli spettri a righe emessi da alcuni elementi chimici. Bohr inizia cosí [1918, p. 1] una sua famosa memoria: « Nel tentativo di sviluppare certe caratteristiche degli spettri a righe, basato su un'opportuna applicazione delle fondamentali idee introdotte da Planck nella sua teoria della radiazione termica alla teoria dell'atomo nucleare di Sir Ernest Rutherford, lo scriven-

te ha mostrato che è possibile in questo modo ottenere una semplice interpretazione di alcune delle piú importanti leggi che governano gli spettri a righe degli elementi...» e dopo aver esposto alcune delle maggiori difficoltà incontrate dalle teorie precedenti aggiunge: «Queste difficoltà sono intimamente connesse con lo stacco radicale dalle idee usuali della meccanica e dell'elettrodinamica implicato nei principî fondamentali della teoria dei quanti [il termine *quantum theory* veniva ormai usato correntemente] e con il fatto che non è stato finora possibile sostituire queste idee con altre che costituiscano una struttura altrettanto sviluppata e consistente». Poche righe piú avanti, all'inizio della parte prima della memoria, vengono enunciati i «principî generali» della nuova teoria: «La teoria quantistica degli spettri di righe riposa sulle seguenti fondamentali ipotesi: I) Un sistema atomico può, e può soltanto, esistere permanentemente in una certa serie di stati corrispondenti a una serie discontinua di valori della sua energia, e conseguentemente ogni cambiamento di energia del sistema, comprendendo fenomeni di emissione e assorbimento di radiazione elettromagnetica, deve aver luogo come una transizione completa tra due tali stati. Questi stati verranno detti gli 'stati stazionari' del sistema. II) La radiazione assorbita o emessa durante una transizione tra due stati stazionari è 'monocromatica' e possiede una frequenza data dalla relazione $E' - E'' = h\nu$, dove h è la costante di Planck e dove E' ed E'' sono i valori dell'energia nei due stati considerati».

Venivano cosí gettate le basi sulle quali si sarebbe dovuta costruire la comprensione di tutti gli spettri atomici e si enunciava contemporaneamente un principio costitutivo delle strutture atomiche per arrivare alla cui chiarificazione e giustificazione un'intera teoria avrebbe dovuto venir costruita. Si sarebbe dovuto cioè edificare quella struttura «altrettanto sviluppata e consistente» che sola avrebbe potuto sostituire l'edificio complesso e per lungo tempo assai soddisfacente della teoria classica.

Per introdurre l'ultimo ingrediente che manca si farà ora un salto indietro fino al 1905, l'anno in cui un ventiseienne impiegato dell'ufficio brevetti di Berna pubblicò sugli «Annalen der Physik» tre memorie tra loro indipendenti e ciascuna delle quali sarebbe bastata a fare di lui Albert Einstein. Si trattava della memoria in cui veniva esposta la teoria della relatività speciale, di quella sul moto browniano e di quella [Einstein 1905], che qui piú interessa, che introduce il concetto di quanto di luce; in essa Einstein postulava per la radiazione elettromagnetica stessa una fondamentale discontinuità; egli in realtà dimostrò, in base a considerazioni in gran parte ancora di natura termodinamica, che anche la radiazione si comporta come se fosse composta di tante piccole unità di radiazione; egli le chiamò «quanti di energia», anche se vennero presto chiamate quanti di luce: per essi Lewis [1926] coniò il termine 'fotoni'. L'importanza dell'osservazione di Einstein – almeno paragonabile a quella delle memorie di Planck del 1900 – era tanto maggiore in quanto sembrava scalzare una delle apparentemente meglio stabilite conquiste della scienza di fine Ottocento, e cioè quella della natura ondulatoria della luce; dopo secoli di *querelle*, con Newton e Laplace da una parte e Huygens, Young e Fresnel dall'altra, la teoria dell'elettromagnetismo di Maxwell sembrava aver stabilito in maniera definitiva che la luce era un

fenomeno ondulatorio, come tale presentava tutti i comportamenti caratteristici di un tale fenomeno (diffrazione e interferenza) e in particolare aveva quindi una struttura tipicamente continua. Le nozioni di onda e di particella, o di corpuscolo, dalla fisica classica cosí accuratamente contrapposte come paradigmi di comportamenti fra loro contraddittori, stavano per venire rimescolate e questa volta in maniera inestricabile; certamente il primo passo in questa direzione fu compiuto da Einstein, anche se le prime evidenze sperimentali dell'effetto fotoelettrico risalgono addirittura allo stesso Hertz (il cui fine era sempre stato al contrario la conferma della teoria di Maxwell!) Nel 1917 Einstein compie un altro passo: parte dall'ipotesi che una molecola di un fissato tipo di gas possa esistere soltanto in un numero discreto di stati; usa la legge di spostamento di Wien (valida, come è noto, soltanto sulla base di considerazioni puramente termodinamiche), oltre a semplici leggi probabilistiche, per ricavare la densità di radiazione in condizioni di equilibrio tra radiazione e molecole: da ciò ricava la legge di distribuzione di Planck e la legge di Bohr che connette la radiazione emessa o assorbita con le energie degli stati iniziale e finale [cfr. Einstein 1917]. In questo modo la nascente «teoria dei quanti» cominciava ad assumere una struttura tale da permettere ai fisici dell'epoca di sperare di proseguire sulla strada imboccata per cercare di costruire quella «struttura altrettanto sviluppata e consistente», da sostituirsi alla meccanica classica, che Bohr invocava nel 1918.

Si è voluto accennare all'opera di Planck, Bohr e Einstein, poiché certamente si tratta dei tre personaggi piú significativi che hanno portato alle decisive elaborazioni degli anni '20. Si accennerà ancora, data la sua importanza anche per la successiva storia della teoria quantistica, a quel principio euristico, o *principio guida*, che fu chiamato da Bohr *principio di corrispondenza* (*Korrespondenzprinzip*). Per comprendere da quale tipo di necessità logica e psicologica un tale principio ebbe origine, occorre ancora una volta ricordare che ormai molti pezzi del composito mosaico della fisica classica ottocentesca stavano venendo scalzati e sconvolti. Il principio di Bohr sulla costituzione dell'atomo era del tutto privo di giustificazioni dal punto di vista classico e anzi una delle ragioni che lo avevano imposto era quella di evitare, per il momento con una ipotesi del tutto ad hoc, le conseguenze disastrose che l'elettromagnetismo classico avrebbe avuto per la stabilità dell'atomo di Bohr: infatti, com'è noto, un tale atomo era un oggetto – classicamente – altamente instabile. Si dirà che questo doveva fornire un buon motivo per abbandonare il modello atomico di Bohr e sostituirlo con uno piú giustificabile, in base, si noti, all'unica teoria all'epoca disponibile e ben stabilita e confermata; ma cosí, come si è visto, non fu: in pratica si giustificò il nuovo modello asserendo che funzionava per definizione, cioè con i nuovi assiomi di Bohr. Come ormai è chiaro, la razionalità interna che percorre l'evolversi dell'operare scientifico è piú sottile di quella che i razionalisti critici vorrebbero scandire nei loro cataloghi normativi e si serve spesso di strumenti che in tali cataloghi vengono bollati di flagrante irrazionalità. Ipotesi ad hoc, contraddizioni e rovesciamenti di posizione apparentemente immotivati sono ben esemplificati nella storia della «old quantum theory».

A proposito delle nuove leggi che tuttavia incominciavano ad emergere nel

nuovo spirito «quantistico», occorre osservare che erano sempre presenti situazioni limite nelle quali in verità le leggi della vecchia fisica classica funzionavano egregiamente. Si ricorderà che la legge di Planck sulla distribuzione delle frequenze irraggiate dal corpo nero nel limite di basse frequenze coincide con la legge di Rayleigh e Jeans, giustificata appunto da ragionamenti classici; una tale caratteristica, lo scostarsi dalla legge classica là dove questa non funziona piú ma il coincidere (con ottima approssimazione) con essa là dove la sua validità era ben confermata, costituisce assai spesso un nodo fondamentale nell'evoluzione delle teorie e fu anche questa volta il motivo ispiratore del principio di corrispondenza. Un tale principio, che ebbe nelle prime memorie di Bohr una formulazione vaga e poi sempre piú precisa [cfr. ad esempio Bohr 1918], asseriva che la nuova teoria quantistica avrebbe dovuto coincidere con la teoria classica nel limite di alti numeri quantici, ovvero per livelli d'energia abbastanza «alti» rispetto al livello piú basso o «fondamentale». In realtà nelle prime memorie di Bohr una tale richiesta sembra venir formulata piuttosto nel limite di basse frequenze (sulla base delle caratteristiche ora ricordate della legge di Planck), e in molti casi interessanti quest'ultimo coincide con il primo pur di intendere il salto energetico tra un livello e il successivo, ma vi sono anche casi in cui ciò non accade, e tra questi perfino l'oscillatore armonico. È, questa dell'insistenza sul principio di corrispondenza, un'esigenza che è stata talvolta mal interpretata; essa non esprime la necessità che eseguendo opportune procedure di limite o simili la nuova meccanica debba coincidere con la teoria classica; si tratta di teorie che poggiano, come si vedrà meglio nel seguito, su concetti interamente differenti e spesso non conciliabili l'uno con l'altro; quello che il principio di corrispondenza richiede è che nel limite di alti numeri quantici vi siano delle coincidenze numeriche tra la nuova e la vecchia teoria; Bohr stesso ha piú volte sottolineato questo aspetto: «la connessione asintotica» tra la nuova meccanica e la fisica classica «qual è supposta nel principio di corrispondenza... non comporta una graduale scomparsa della differenza tra il trattamento quantistico dei fenomeni della radiazione e le idee dell'elettrodinamica classica; tutto ciò che si afferma è un accordo asintotico dei risultati numerici statistici» [1922, p. 144].

Va peraltro osservato che il potere di resistenza dell'elettrodinamica classica fu molto alto per tutto il primo quarto del nostro secolo; anche nell'ambito della descrizione del comportamento dei corpi microscopici, dove erano sembrate inequivocabili le sue cadute, per molti anni si tese ad usarla per quel che poteva ancora dare. Fu soltanto con l'avvento della teoria quantistica dei campi, ed in particolare dell'elettrodinamica quantistica, che si ebbe in mano qualcosa di alternativo con cui sostituirla.

4. *La meccanica quantistica prende forma.*

Negli anni che precedettero immediatamente il 1925 avvennero alcuni sviluppi cruciali per il futuro della meccanica quantistica ed indicativi del livello di astrattezza da cui questa doveva essere caratterizzata.

Anzitutto comparvero nella descrizione dei fenomeni fisici di cui ci si andava interessando, interazione fra atomi e radiazione, entità qualificate dall'aggettivo 'virtuale', a significare che tali entità non erano proprio ciò che si voleva descrivere, ma si comportavano *come se* fossero state ciò che si voleva descrivere. Si vuole qui alludere anzitutto all'introduzione dei concetti di oscillatore virtuale, già implicata da Ladenburg [1921], e di campo virtuale di radiazione, introdotta da Slater qualche anno piú tardi [1924]: l'idea sottostante a tale introduzione era che un atomo poteva venir descritto come un insieme di oscillatori, detti appunto virtuali (termine coniato probabilmente da Bohr), con frequenze proprie pari alle frequenze di assorbimento di quell'atomo; s'intende che le frequenze d'assorbimento di un atomo sono quelle tali che l'atomo possa assorbire, eseguendo un passaggio di stato (eccitazione da uno stato inferiore a uno stato superiore), un *quanto* di radiazione elettromagnetica caratterizzata da tale frequenza; l'idea di campo virtuale di radiazione riguardava il campo generato appunto dagli oscillatori che simulavano il comportamento del singolo atomo; tale campo era l'entità che avrebbe dovuto mediare l'interazione con gli altri atomi; la sua funzione, cosí si esprime Slater [*ibid.*, p. 308] «è quella di assicurare una conservazione statistica dell'energia e del momento determinando le probabilità delle transizioni quantistiche». «La teoria qui esposta, – scrive Rosenfeld [1970, p. 81] alludendo al successivo Bohr, Kramers e Slater [1924], – appare, in retrospettiva, non come un episodio di breve respiro nella storia della fisica atomica, ma come un manifesto destinato ad avere una perdurante influenza, che contiene un programma la cui esecuzione rivelò aspetti essenziali della struttura della materia e della luce». È un'idea dirompente rispetto a tutta la tradizione precedente: quella che l'energia si conservasse in maniera soltanto statistica, non cioè necessariamente in ogni singolo processo elementare di interazione fra atomo e atomo o fra atomo e radiazione, ma soltanto complessivamente, ovvero che le probabilità di transizione o di verificarsi di vari processi elementari fossero tali che la somma di tutte le energie si conservasse; se si pensa a quel che aveva voluto dire il teorema di conservazione dell'energia in tutti i decenni precedenti, anche, si noti, nel processo di distacco dalla teoria classica, questa disponibilità a indebolirlo cosí pesantemente appare abbastanza sorprendente; sembra anzi che Slater stesso non fosse del tutto convinto della cosa ma che si fosse fatto convincere dal continuo dibattito con Kramers e Bohr [Waerden 1967, p. 13].

L'espressione 'meccanica quantistica' (*Quantenmechanik*) appare per la prima volta come titolo di un articolo di Born degli stessi anni (1924). Born – già dal 1921 – era professore ordinario di fisica teorica a Göttingen, sulla cattedra che era stata per breve tempo di Debye, poi trasferitosi a Zurigo, e prima di lui di Voigt. Tra i suoi assistenti contò, in tempi diversi, personaggi che presto sarebbero divenuti di primo piano quali Heisenberg, Jordan, Heitler e Rosenfeld. L'interazione che si verificò negli anni 1924-26 tra Born, Jordan e Heisenberg e tra quest'ultimo e Bohr fu cruciale per il sorgere della nuova teoria; questa prese in un primo tempo il nome di «meccanica delle matrici» per il ruolo decisivo che vi svolgevano le regole di moltiplicazione tra matrici, e quindi tra i loro elementi (le cosiddette «ampiezze di probabilità»); ruolo proposto ed evidenziato

soprattutto da Born [cfr. per questa interazione Bohr 1961; Born 1940-63; Heisenberg 1964].

Colui che eseguí un primo passo decisivo fu Heisenberg. Questi, studente a Monaco del grande Sommerfeld, aveva visitato Göttingen per la prima volta, ventunenne, nel 1922, in occasione di alcune lezioni ivi tenute da Bohr, dette poi «Bohr-Festspiele». Dopo esser tornato a Monaco a scrivere la sua tesi di dottorato, divenne, nell'ottobre 1923, assistente di Born a Göttingen e con questi pubblicò rapidamente alcuni lavori. Cruciale per Heisenberg sembra essere stato già anche l'incontro con Bohr nel 1922. Scrive Heisenberg stesso [1964]: «Quando la discussione fu terminata, Bohr venne da me e propose che si facesse una passeggiata insieme sulle alture di Hainberg, fuori Göttingen. Naturalmente, accettai volentieri. Quella discussione, che ci portò avanti e indietro sulle boscose alture di Hainberg, fu la prima discussione particolareggiata che riesco a ricordare sulla fisica fondamentale e sui problemi filosofici della fisica atomica, ed ha certamente avuto sulla mia futura carriera un influsso decisivo. Capii per la prima volta che il punto di vista di Bohr sulla sua stessa teoria era assai piú scettico di quello di molti altri fisici – ad esempio Sommerfeld – del tempo e che la sua intima comprensione della struttura della teoria non era il risultato di un'analisi matematica delle ipotesi di base, ma un intenso coinvolgimento con la realtà dei fenomeni, cosí che gli era possibile avvertire intuitivamente le relazioni piuttosto che derivarle formalmente. Cosí compresi: la conoscenza della natura veniva ottenuta in prima istanza in questo modo, e solo come passo successivo si può riuscire a fissare la propria conoscenza in forma matematica cosí da assoggettarla a una completa analisi razionale. Bohr era in primo luogo un filosofo, e non un fisico, ma egli si rendeva conto che nella nostra epoca la filosofia naturale acquista rilevanza soltanto se ogni suo particolare può essere sottoposto all'inesorabile vaglio dell'esperienza» (trad. ingl. p. 95). L'interazione con Bohr proseguí in verità assai intensamente anche negli anni successivi, durante lunghi soggiorni del giovane assistente di Born a Copenhagen. Infine, nel giugno del 1925, Heisenberg passò una decina di giorni sulla sperduta isoletta di Helgoland, al largo delle coste danesi nel mare del Nord, per sfuggire alla febbre da fieno che lo aveva colpito sul continente; fu in quei pochi giorni, secondo la sua stessa testimonianza, che uno sprazzo di luce si fece strada in lui fino a permettergli una prima formulazione delle nuove «regole» che dovevano formare la «meccanica delle matrici».

Forse la cosa piú importante che Heisenberg comprese durante quei giorni, e sul cui senso si sentiva «sempre piú radicale», fu il fatto che la meccanica classica non poteva piú essere usata in nessun punto del ragionamento; in verità quel che era avvenuto fino a quel momento era che il calcolo di alcune caratteristiche degli oscillatori virtuali era fatto sfruttando la trattazione classica dell'oscillatore armonico, e dando luogo quindi in verità a una complicata mistura, razionalmente assai poco giustificabile, nella quale alla vecchia teoria classica erano semplicemente giustapposte le regole della nuova *Quantentheorie*. D'altra parte conservava senso tener fede al bohriano principio di corrispondenza secondo il quale, come si è visto, per alti numeri quantici i risultati della nuova teoria dovevano ben approssimare (o coincidere con) quelli della teoria classica.

Heisenberg, riprendendo e concretando un'esigenza che già era stata sottolineata da Poincaré [1912], comprese che non solo nuove regole di calcolo andavano introdotte per le interazioni e i salti quantici dei livelli atomici, ma una nuova cinematica cioè una nuova *descrizione* del moto. In particolare egli propose che la classica equazione del moto $\ddot{x}+f(x)=0$ conservasse validità soltanto al prezzo di cambiare profondamente l'interpretazione della funzione $x(t)$, che non avrebbe piú dovuto rappresentare, come nella meccanica classica, la posizione dell'elettrone in funzione del tempo, bensí qualche nuova grandezza connessa con le probabilità di transizione dell'elettrone tra certe coppie di stati. Una delle motivazioni fondamentali di Heisenberg, e ciò sembra epistemologicamente cruciale, fu il fatto che mentre la traiettoria, cioè la funzione che dà la posizione al variare del tempo, per un elettrone è scarsamente osservabile, le grandezze in prima istanza osservabili sono le probabilità di transizione fra stati. A questo primo lavoro di Heisenberg [1925] (inviato alla «Zeitschrift für Physik» negli ultimi giorni di luglio da Born, cui il suo assistente l'aveva sottoposto per approvazione) seguí un articolo congiunto di Born e Jordan [1925], nel quale viene data l'interpretazione matriciale delle regole di calcolo di Heisenberg, viene proposta e dimostrata la regola di commutazione $pq-qp=(h/2\pi i)1$ e il teorema di conservazione dell'energia, e infine il piú famoso *Dreimännerarbeit* [Born, Heisenberg e Jordan 1925] in cui la nuova meccanica veniva già formulata con notevole completezza.

L'aspetto di fondo, che è opportuno menzionare fin d'ora, per cui la nuova teoria si caratterizzava fortemente rispetto alla fisica classica è l'importanza esclusiva accordata alle quantità «osservabili»: «In questa situazione sembra sensato abbandonare qualsiasi speranza di osservare quantità finora non osservabili, quali la posizione e il periodo di un elettrone, e ammettere che il parziale accordo delle regole quantistiche (le "quantum rules" degli approcci parziali precedenti) con l'esperienza sia piú o meno fortuito. Sembra invece piú ragionevole cercar di stabilire una meccanica teorica quantistica, analoga alla meccanica classica, ma nella quale entrino soltanto relazioni tra quantità osservabili» [Heisenberg 1925, p. 879]. Fu questo un principio guida cruciale per chi contribuí a mettere insieme, negli ultimi anni '20, quella «mistura di congetture interessanti, dichiarazioni dogmatiche ed assurdità filosofiche» [Feyerabend 1968, p. 310] che fu presto compresa nella locuzione, certo un po' sommaria, di «Kopenhagener Deutung»: l'interpretazione di Copenhagen, che, pur non essendo priva di alternative, è stata ed è dominante nella pratica corrente dei fisici che di meccanica quantistica tuttora si occupano come strumento di lavoro. Un principio su cui si avrà modo di tornare quando ci si sarà soffermati un po' di piú sulla struttura della teoria, che implicava un atteggiamento abbastanza nuovo rispetto alle aspettative dello scienziato e del ricercatore: meno conta il render conto in maniera esaustiva di quelle proprietà che, intuitivamente (*ovvero* per analogia con la meccanica classica), si sarebbe portati ad attribuire ad ogni sistema fisico, semplice o complesso, e piú conta esaminare attentamente e spregiudicatamente quali sono le grandezze fisiche osservabili di un determinato sistema – quelle che sul serio si possono misurare perché si possono immaginare almeno in linea di

principio degli strumenti in grado di farlo – e costruire una teoria che di quelle parli e che di quelle studi e predíca i valori. Meno spiegazione di meccanismi «intimi» e forse mai osservabili e piú descrizione accurata dei risultati che effettivamente è dato osservare. Come si vedrà, questa distinzione su scala atomica è assai importante, anche se naturalmente non è facile raggiungere l'unanimità quando si deve decidere che cosa è in linea di principio osservabile e che cosa non lo è.

Su tali basi, con il contributo essenziale, oltre che di Born, Heisenberg e Jordan, di Bohr, Pauli ed altri, prese forma molto rapidamente la nuova teoria, che sostituiva la meccanica classica per trattare il comportamento di sistemi meccanici atomici e che è ancora oggi in qualche misura dominante e di uso continuo in fisica dei sistemi microscopici: essa è nota come meccanica quantistica. Oltre ai contributi cui si è accennato non può essere dimenticato l'apporto fondamentale dell'inglese Dirac, allora giovane fisico a Cambridge, e quindi leggermente defilato rispetto all'ambiente assai creativo della fisica tedesca e di Copenhagen. Nel luglio del 1925 Heisenberg si recò per una serie di lezioni a Cambridge e Dirac ne ebbe notizia, anche se in maniera molto indiretta e frammentaria. Nei mesi successivi il fisico inglese elaborò, talvolta riscoprendo, ma spesso aggiungendo e completando, gran parte della teoria, che espose nello stesso anno [Dirac 1926*a*] e che generalizzò notevolmente l'anno successivo [Dirac 1926*b*], applicando sistematicamente e con successo il suo formalismo alla trattazione dell'atomo d'idrogeno.

5. *L'alternativa ondulatoria.*

Anche per la «meccanica ondulatoria» che per qualche tempo parve rappresentare un'alternativa globale alla formulazione sviluppata a Göttingen e a Copenhagen, il punto di partenza può essere ricondotto alle memorie scritte da Planck nel 1900; personaggi chiave per il suo evolversi furono essenzialmente in primo luogo Broglie (cfr. in particolare la sua già citata tesi di dottorato); in secondo luogo lo stesso Einstein, che non nascondeva una notevole insoddisfazione nei confronti del punto di vista, per alcuni aspetti definibile «astratto», che andava emergendo a Göttingen e a Copenhagen, e infine Schrödinger. Si ricorderà qui che l'ipotesi basilare cui era giunto Louis de Broglie, a partire da considerazioni che implicavano la legge di trasformazione delle frequenze – secondo la teoria della relatività ristretta – al variare del sistema di riferimento, consisteva nel fatto che ad ogni particella in moto deve essere associata un'onda, di frequenza ed ampiezza opportune, e che fosse impossibile separare il moto della particella e la propagazione dell'onda; in particolare Broglie dimostrò che la velocità della particella doveva coincidere con la velocità di gruppo delle onde di fase. Su questa via egli arrivò ad osservare che le condizioni di quantizzazione di Bohr-Sommerfeld potevano venir lette come una prescrizione riguardante il numero di lunghezze d'onda necessarie a sovrapporsi esattamente all'orbita di un elettrone intorno al nucleo. In altre parole, dalla condizione di Bohr-Sommerfeld $\int p\,dq = nh$ e dalla relazione di Broglie $p = h/\lambda$ fra il momento p dell'elettrone e la lunghezza d'onda λ ad esso associata seguiva una condizione del tipo $\int(1/\lambda)\,dq = n$. A Schrödinger, che per vari anni addietro s'era occupato, in vari ambiti della fisica classica, di problemi agli autovalori, una tale equazione suonò immediatamente significativa. Nella sua prima memoria sull'argomento [1926*a*], egli propose quella che è ora nota come equazione di Schrödinger indipendente dal tempo. Sostituendo nell'equazione di Hamilton-Jacobi $H(q,(\partial S/\partial q)) = E$ al posto della funzione S (assunta separabile) l'espressione $K \ln\psi$, dove ψ, dipendente dalla variabile di posizione q, viene ora interpretata, anche qui sulla base di un'analogia proveniente dall'ottica classica, come «funzione d'onda», Schrödinger fu in grado di esprimere la condizione che doveva individuare i valori dell'energia permessi, imponendo che il funzionale corrispondente avesse un estremo in corrispondenza di tali valori. L'equazione di Eulero-Lagrange risultante da un tale problema variazionale è l'«equazione d'onda».

Lo sforzo di Schrödinger si concentrò nella direzione di fornire per l'interpretazione fisica, non meno che per la trattazione matematica della funzione d'onda ψ, degli argomenti che riposassero, e ciò in netta antitesi con lo «spirito di Copenhagen», in ultima analisi su una modellizzazione fornita di una sicura base meccanico-intuitiva. In altre parole può ben dirsi che Schrödinger, assai meno di Heisenberg, era disposto ad adottare quella nuova cinematica che avrebbe formato la caratteristica dominante della meccanica di Copenhagen, e che la sua maggior preoccupazione consisteva nel non perdere il contatto con l'intuizione fisica. Vicino a tale posizione rimase fino alla sua morte lo stesso Einstein; di notevole interesse è a questo proposito la lettura del carteggio tra i due [Schrödinger e altri 1926-50, pp. 21-37].

Nel 1926 Schrödinger portò comunque a termine l'edificazione della sua teoria, giungendo a scrivere non solo un'equazione che determinasse i valori dell'energia permessi per un dato sistema atomico, ma anche un'equazione che fornisse la dinamica di tale sistema, il che voleva dire l'evoluzione nel tempo della sua funzione d'onda. Schrödinger riuscí anche a mostrare, almeno in un certo senso, l'*equivalenza* tra i propri risultati e quelli, da poco pubblicati, di Heisenberg, Born e Jordan. Si riporta qui, per il suo notevole interesse, il primo capoverso di tale memoria [1926*b*, p. 734]: «Se si considera la straordinaria divergenza di punti di partenza e di concetti tra la meccanica quantistica di Heisenberg da un lato e la teoria recentemente esposta in questa sede nei suoi lineamenti fondamentali e chiamata meccanica "ondulatoria" o "fisica" dall'altro, è molto strano che, per quanto riguarda i risultati particolari resi noti finora, queste due nuove teorie quantistiche *siano in accordo tra loro* anche laddove esse differiscono dalla vecchia teoria quantistica. Menziono in particolare la peculiarità dei numeri "semi-interi" che appaiono nell'oscillatore e nel rotatore. Ciò è in realtà assai notevole poiché punto di partenza, concetti, metodi e l'intero apparato matematico appaiono di fatto essere diversi. Soprattutto, tuttavia, lo stacco dalla meccanica classica appare avvenire nelle due teorie in direzioni che sono in realtà diametralmente opposte. Heisenberg sostituisce le variabili continue classiche con insiemi di quantità numeriche discrete (matrici), che dipendono

da una coppia di indici interi e che sono determinate da equazioni *algebriche*. Gli stessi autori chiamano la teoria una "vera teoria del discontinuo". Sull'altro lato la meccanica delle onde implica – esattamente al contrario – un passo dalla meccanica classica verso *la teoria del continuo*. Appare infatti, al posto del fenomeno descrivibile mediante un numero finito di variabili dipendenti con un numero finito di equazioni differenziali ordinarie, un fenomeno di *campo* continuo nello spazio delle configurazioni che è governato da una singola equazione alle derivate *parziali* derivabile da un principio d'azione. Questo principio d'azione, o quest'equazione differenziale, sostituisce le equazioni del moto *e* le condizioni quantistiche della precedente "teoria quantistica classica" ».

I dieci anni successivi al citato testo di Schrödinger videro la sistemazione e la stabilizzazione della teoria [cfr. Sommerfeld 1929; Born e Jordan 1930; Landé 1928; Pauli 1933], e soprattutto i fondamentali *Principles* di Dirac [1930], seguiti dalla sistemazione dotata di grande rigore formale di Neumann [1932]: è in quest'ultima forma che si forniranno qui gli elementi di base della nuova teoria.

6. *La struttura formale della nuova meccanica.*

6.1. Cinematica quantistica.

Lo strumento matematico cardine che venne definitivamente adottato per una formulazione adeguata della meccanica quantistica fu quello della teoria degli operatori lineari in uno spazio di Hilbert e dei loro autovalori e autovettori.

S'è detto «definitivamente» affinché non si pensi che sia stato questo il primo e piú naturale ambito formale nel quale ci si sia mossi per dare una veste rigorosa alla nuova teoria. Come le parole di Schrödinger dovrebbero aver suggerito, da un lato le regole di calcolo proposte dai fisici di Göttingen richiedevano la manipolazione di equazioni algebriche (spesso infinite) con (corrispondentemente infinite) incognite: tali incognite erano le cosiddette «ampiezze di probabilità» i cui (moduli) quadrati dovevano fornire le probabilità di transizione fra diversi stati quantici. D'altro lato la linea dello stesso Schrödinger richiedeva in maniera naturale l'uso di equazioni differenziali, la teoria delle loro singolarità e la ricerca dei loro autovalori. È in verità questa seconda l'impostazione che si rivelò da molti punti di vista piú maneggevole, che venne piú spesso usata nelle presentazioni offerte dai primi manuali e in ultima analisi la piú suscettibile di una trattazione rigorosa.

Si osserverà anche che la teoria degli operatori lineari in spazi di Hilbert, già in parte disponibile a quell'epoca, ricevette grande impulso dalle crescenti richieste provenienti dalla fisica teorica, cosí che anche il suo attuale e assai notevole sviluppo può dirsi in gran parte stimolato dalle molteplici necessità di quest'ultima.

Qui è opportuno introdurre un esempio di spazio di Hilbert che riveste una importanza cruciale nei casi piú semplici di sistema fisico trattato quantisticamente, quello dello spazio $L^2(\mathbf{R})$ (cfr. anche l'articolo «Calcolo» in questa stessa *Enciclopedia*, II, pp. 444-45); s'intende con questo simbolo l'insieme costituito da tutte le funzioni definite sulla retta reale $\mathbf{R}$ e a valori in $\mathbf{C}$, che siano misurabili e a modulo quadrato integrabile; ciò significa che per una di tali funzioni, detta f, esiste finito l'integrale (qui e nel seguito definito rispetto alla misura m di Lebesgue) $\int_{\mathbf{R}}|f|^2\,dm$, ovviamente positivo o nullo; cruciale è la circostanza che non si chiede a tali funzioni di essere continue ma solo, molto piú debolmente, di essere misurabili, e ciò essenzialmente perché abbia senso parlare dell'integrale; la struttura lineare di tale insieme si ottiene facilmente ricorrendo alla somma puntuale di funzioni $((f+g)(x):=f(x)+g(x))$ e all'usuale prodotto per gli scalari di $\mathbf{C}((\alpha f)(x):=\alpha f(x))$; il prodotto interno è definito dalla formula: $(f|g):=\int_{\mathbf{R}}\bar{f}g\,dm$; in verità cosí facendo s'incorre nella difficoltà che le funzioni quasi ovunque nulle su $\mathbf{R}$ avrebbero norma nulla pur non coincidendo con la funzione identicamente uguale a zero; a ciò si ovvia facilmente considerando, anziché le funzioni vere e proprie, classi di equivalenza di funzioni, ponendo come relazione di equivalenza l'uguaglianza quasi ovunque su $\mathbf{R}$, sempre rispetto alla misura di Lebesgue; qui e nel seguito, ove la cosa non generi ambiguità, si prescinderà sempre da tale distinzione e si parlerà tranquillamente di funzioni anche se la dizione rigorosa sarebbe «classi di equivalenza di funzioni»; si ha anche che la norma di un singolo elemento $f\in L^2(\mathbf{R})$ è data dall'espressione $\sqrt{\int|f|^2\,dm}$; opportuni ragionamenti e dimostrazioni mostrerebbero ora che l'insieme cosí costruito, con la struttura in esso definita, è uno spazio di Hilbert. Il significato di questo particolare spazio di Hilbert per la meccanica quantistica è il seguente: si supponga di voler descrivere un sistema fisico tra i piú semplici pensabili, una particella puntiforme vincolata a muoversi su una retta, o, come talvolta si dice, «la particella in una dimensione». La meccanica quantistica, già dal punto di vista cinematico, è una teoria profondamente diversa dalle teorie classiche; non descrive la particella dicendo qual è la sua traiettoria, cioè dicendo in ogni istante qual è la sua posizione; un tale modo di fare sarebbe incompatibile con i principî della nuova teoria; la descrizione piú completa possibile della situazione, o, come si dice, dello *stato* della particella, consiste nell'assegnare una funzione ψ (appunto un elemento di $L^2(\mathbf{R})$) che rappresenti per la particella la seguente informazione: fissato un qualunque intervallo K contenuto in $\mathbf{R}$, la probabilità che una misura della posizione della particella dia come risultato un punto appartenente all'intervallo K è data da $\int_K|\psi|^2\,dm$; ciò viene talvolta espresso, in maniera un po' meno precisa, dicendo che la probabilità di trovare la particella nell'intervallo $(x, x+dx)$ è data dalla quantità $|\psi(x)|^2\,dx$; si noti che da ciò segue ad esempio che se ci si chiedesse quanto vale la probabilità di trovare la particella esattamente nel punto x la risposta sarebbe zero, poiché l'integrale di una qualunque funzione di $L^2(\mathbf{R})$ esteso ad un solo punto di $\mathbf{R}$ è comunque zero; ciò corrisponde al fatto, sperimentalmente assai ragionevole, che una misura di posizione non determina mai con esattezza *un punto* della retta, ma un intervallo, eventualmente assai piccolo, nel quale la particella è situata. L'aspetto cruciale che differenzia una tale descrizione da quella cui si era abituati dalla teoria classica non è evidentemente il fatto che una tale funzione possa essere asse-

gnata (anche per una particella che esegua una traiettoria nel senso classico del termine sarebbe perfettamente possibile calcolare una tale funzione), ma il fatto che in questo caso è questa *la descrizione piú completa possibile*, e ciò non per ragioni sperimentali in linea di principio superabili, ma per caratteristiche intrinseche della nuova teoria che non prevede, proprio in linea di principio, descrizioni piú dettagliate di questa. Si noti d'altra parte che una funzione nulla al di fuori di un intervallo di ampiezza piccola a piacere, ad esempio δ, e costante all'interno di tale intervallo, appartiene certamente a $L^2(\mathbf{R})$ ed è dunque possibile, anche all'interno della nuova meccanica, descrivere una particella confinata in un intervallo determinato con precisione arbitraria; tuttavia la differenza formale rispetto alla descrizione classica è assai rilevante ed implica, come si vedrà, conseguenze di grande portata.

Quanto detto finora riguarda la descrizione cinematica del moto di una particella e potrebbe essere facilmente esteso, pur d'introdurre complicazioni concettualmente non essenziali, a sistemi piú complicati che non la particella in una dimensione. In generale cioè uno stato di un sistema fisico viene descritto da un vettore di un opportuno spazio di Hilbert, o, meglio, da un vettore normalizzato (la necessità di una tale normalizzazione proviene dall'interpretazione probabilistica già accennata e sulla quale si tornerà).

Rimane ancora da descrivere come vengono rappresentate, in questo formalismo che sembra assai meno «intuitivo» di quello cui la meccanica classica ci aveva abituato, le «osservabili», cioè le grandezze fisiche che caratterizzano il sistema, ad esempio il momento lineare, l'energia, il momento angolare ed eventuali altre proprie della nuova teoria. Per far ciò occorre brevemente ricordare che un operatore lineare in uno spazio di Hilbert $\mathcal{H}$ è una funzione, definita su un dominio che sia una varietà lineare dello spazio e a valori sempre nello spazio in questione, con le proprietà $A(x+y)=Ax+Ay$ e $A(\alpha x)=\alpha Ax$, per ogni possibile coppia di vettori x e y di $\mathcal{H}$ e per ogni numero complesso α. È opportuno ricordare che la classe di operatori che hanno proprietà «adatte» a far sí che possano rappresentare osservabili quantistiche è quella degli operatori cosiddetti autoaggiunti. Si veda ad esempio questa nozione nell'articolo «Infinitesimale» di questa stessa *Enciclopedia*, VII, p. 516. Si ha, per tali operatori: $(Ax|y)=(x|Ay)$; da cui si desume anche, per i vettori del dominio di A, che $(x|Ax)$ è un numero reale (e non semplicemente un generico numero complesso); questa circostanza permetterà l'interpretazione fisica voluta per tali numeri. Vi è però un'altra caratteristica assai importante degli operatori autoaggiunti, di meno facile enunciazione ma di significato assai piú profondo: si tratta del cosiddetto teorema spettrale. La problematica relativa all'estensione della classica teoria spettrale (per spazi a dimensione finita) a operatori definiti sugli spazi di Hilbert può ad esempio essere vista in «Infinitesimale», pp. 513-15.

Nell'interpretazione fisica un operatore autoaggiunto rappresenta un'osservabile; quando esso possiede degli autovettori con relativi autovalori, questi rappresentano tutti e soli i risultati che possono essere ottenuti in misure di questa osservabile; se il sistema fisico considerato si trova in uno stato rappresentato da un vettore x, appartenente al dominio dell'operatore, per sapere quale risultato

si otterrà misurando l'osservabile rappresentata da A su tale stato, occorre scrivere x in termini del sistema completo di autovettori di A, $x=\sum_i w_i x_i$; la probabilità di ottenere nella misura il valore a_i, autovalore dell'autovettore x_i, è data dal quadrato del modulo del prodotto interno $(x_i|x)$, che, in questo caso, è dato da $|w_i|^2$; questo è tutto quanto si può dire a proposito dei risultati possibili di misure dell'osservabile A sul sistema nello stato x. Vi saranno casi particolarmente fortunati: ad esempio potrebbe essere che x coincida con uno degli autovettori x_i; sia $x=x_{i_0}$; evidentemente è allora $w_{i_0}=1$ e $w_i=0$ se $i\neq i_0$; in questo caso vi è quindi probabilità nulla di trovare un valore diverso da a_{i_0} e vi è probabilità pari a 1 di trovare il valore a_{i_0}; si è quindi certi di trovare tale valore; a parte questo caso, come si vede, non vi può essere certezza, ma solo probabilità. Si osserverà che il fatto di aver scelto di rappresentare i sistemi con vettori *normalizzati* dello spazio di Hilbert fa sí che si abbia in ogni caso $\sum_i |w_i|^2=1$, cosí che la probabilità totale di trovare *qualche* valore nella misura di A sia effettivamente unitaria. Ora il teorema spettrale nella sua piú generale enunciazione permette appunto di estendere affermazioni di questo tipo ad operatori autoaggiunti che non posseggono un sistema completo di autovettori; tale teorema permette di associare ad ogni operatore autoaggiunto una famiglia di operatori di proiezione, o proiettori: ad ogni intervallo K di $\mathbf{R}$ (in verità anche a sottoinsiemi assai piú generali di $\mathbf{R}$) è associato un proiettore P_K; esso è univocamente determinato dall'operatore autoaggiunto A e dall'intervallo K; la probabilità di trovare, in una misura dell'osservabile A sul sistema che si trova nello stato x, un valore compreso nell'intervallo K è data da $(x|P_K x)$. La coerenza dell'interpretazione probabilistica è qui assicurata dal fatto che il teorema spettrale garantisce che il proiettore associato all'intero asse reale $P_{\mathbf{R}}$ è, qualunque sia l'operatore A, sempre il proiettore unità 1, cosí che $(x|P_{\mathbf{R}}x)=(x|x)=1$. Un caso particolare di questa situazione si ha nell'esempio ricordato inizialmente per dare il significato della funzione ψ per la particella in una dimensione; l'operatore Q che rappresenta l'osservabile posizione in questo caso è dato da $(Q\psi)(t)=t\,\psi(t)$, su un opportuno dominio di funzioni di $L^2(\mathbf{R})$. Il teorema spettrale associa a tale operatore, autoaggiunto in $L^2(\mathbf{R})$, e che non possiede né autovalori né autovettori nel suo dominio, la famiglia di proiettori definiti da $(P_K\psi)(t)=\chi_K(t)\,\psi(t)$ ove si è indicata con χ_K la funzione caratteristica dell'insieme K (la funzione caratteristica di un sottoinsieme di un insieme dato vale 1 nei punti del sottoinsieme e vale 0 nei punti del suo complementare): ne segue che la probabilità di trovare, facendo una misura dell'osservabile posizione, un valore compreso in K, è data da $(\psi|P_K\psi)=\int_{\mathbf{R}}\bar{\psi}(t)\,\chi_K(t)\,\psi(t)\,dt=\int_K|\psi(t)|^2\,dt$ in accordo con quanto visto piú sopra.

6.2. Misura e indeterminazione.

Se un sistema fisico si trova nello stato rappresentato da un autovettore di un operatore autoaggiunto relativo a una determinata osservabile, una misura di tale osservabile sul sistema dà come risultato necessariamente l'autovalore corrispon-

dente; è vero anche che, se facendo una misura sul sistema di quella osservabile si trova con certezza un dato autovalore come risultato, allora il sistema si trova nell'autostato (= stato rappresentato dall'autovettore) corrispondente; da ciò viene dedotta la seguente conseguenza, e viene dedotta, si noti, basandosi su un principio di «continuità fisica» [Dirac 1930, trad. it. p. 48] che sembrava non avrebbe piú dovuto aver diritto di cittadinanza nella nuova meccanica: «Quando effettuiamo la misura di una variabile dinamica reale ξ, la perturbazione che l'atto della misura comporta provoca una brusca variazione nello stato del sistema dinamico. Per continuità fisica, se immediatamente dopo la prima misura ne facciamo una seconda della stessa variabile dinamica ξ, il risultato di quest'ultima deve coincidere con quello della prima. Dunque, una volta eseguita la prima misura, non vi è piú indeterminazione nel risultato della seconda. Pertanto, dopo la prima misura, il sistema si trova in un autostato della variabile dinamica ξ appartenente allo stesso autovalore della prima. Tale conclusione deve continuare a valere anche se la seconda misura non viene in pratica eseguita». Ciò implica che l'esecuzione di una misura provoca un brusco cambiamento nello stato del sistema, almeno se questo già non si trovava in un autostato dell'osservabile di cui si tratta. È questo uno dei tratti fondamentali della meccanica quantistica, su cui si avrà modo di tornare, e che la distingue ulteriormente in maniera molto pesante, dal punto di vista della teoria della conoscenza, dalla teoria classica. La misura interferisce con il sistema fisico su cui viene eseguita, non è un'osservazione passiva di qualcosa di esistente ed evolventesi in maniera totalmente indipendente da essa. Chi per primo teorizzò in modo sistematico tale punto di vista, sia dal punto di vista formale, sia, cosa assai piú significativa, dal punto di vista di una analisi delle effettive procedure sperimentali che portano – su scala atomica – a «leggere» o meno determinati valori di determinate osservabili fu, come già s'è accennato, Heisenberg. «Nelle teorie classiche, – cosí egli si esprime [1930, trad. it. p. 13], – quest'interazione [fra oggetto e osservatore] veniva considerata o come trascurabilmente piccola o come controllabile, in modo tale da poterne eliminare in seguito l'influenza, per mezzo di calcoli. Nella fisica atomica, invece, tale ammissione non si può fare, poiché a causa della discontinuità degli avvenimenti atomici ogni interazione può produrre variazioni parzialmente incontrollabili e relativamente grandi. Questa circostanza ha come conseguenza il fatto che in generale le esperienze eseguite per determinare una grandezza fisica rendono illusoria la conoscenza di altre grandezze ottenute precedentemente; esse infatti influenzano il sistema su cui si opera in modo incontrollabile e quindi i valori delle grandezze precedentemente conosciute ne risultano alterati. Se si tratta questa perturbazione in modo quantitativo, si trova che in molti casi esiste, per la conoscenza contemporanea di diverse variabili, un limite di esattezza finito, che non può essere superato».

Le ultime frasi introducono un tema di grande rilevanza per la teoria quantistica, quello cioè delle osservabili *compatibili* o no. Si allude, anche se in maniera soltanto qualitativa, al fatto che, poiché una misura di un'osservabile perturba in qualche modo il sistema, una misura di qualche altra osservabile può rendersi del tutto ineffettuabile. Un'analisi piú particolareggiata mostra che in verità

vi sono osservabili tra loro «incompatibili» ed altre che invece sono compatibili, tali cioè da poter essere misurate simultaneamente su uno stesso sistema. L'esempio piú tipico di osservabili incompatibili è offerto dalla posizione e dal momento lineare; limitandosi al caso della particella in una dimensione, si è visto quale forma assume l'operatore che rappresenta l'osservabile posizione; l'osservabile momento lineare della particella (si ricordi che il momento lineare è classicamente un vettore dato dal prodotto della massa per la velocità) è rappresentata dall'operatore, spesso denotato con P, definito, su un opportuno dominio in $L^2(\mathbf{R})$, dalla formula $Pf = -i\hbar f'$, dove l'apice sta ad indicare l'operazione di derivazione (nel dominio di P vi sono soltanto funzioni per cui tale operazione ha senso); la ricordata non-compatibilità tra le osservabili posizione e momento lineare è espressa, nel formalismo che descrive la teoria, dal fatto che gli operatori, entrambi autoaggiunti, Q e P *non commutano* tra loro; ciò significa che se si applicano a un elemento di $L^2(\mathbf{R})$ per cui ciò ha senso, prima l'operatore QP e poi l'operatore PQ (il prodotto di due operatori si esegue applicando prima l'operatore di destra e poi quello di sinistra) si ottengono due risultati diversi; un semplice calcolo mostra infatti che il commutatore di Q e P, cioè l'espressione $QP-PQ$ usualmente denotata $[Q,P]$ non è nullo e agisce come un multiplo dell'operatore unità: $[Q,P]f = i\hbar f$. Questa situazione è molto generale; in ogni sistema fisico, ad esempio formato da molte particelle, per il quale sia definita una posizione, o meglio, visto che si considerano in generale sistemi realistici in tre dimensioni, le componenti della posizione, e le componenti del momento lineare, le relazioni di commutazione tra tali componenti sono $[Q_x,P_x] = i\hbar$, $[Q_y,P_y] = i\hbar$, $[Q_z,P_z] = i\hbar$, mentre tutte le altre (ad esempio $[Q_x,P_y]$ ecc.) sono nulle: componenti diverse della posizione e del momento lineare sono tra loro compatibili, esse possono venir contemporaneamente misurate; le componenti dello stesso tipo sono invece incompatibili. Vi è una semplice giustificazione formale di questa incompatibilità di osservabili tra loro non commutanti (cioè il cui commutatore non è nullo); dagli elementi già ricordati a proposito della cinematica probabilistica che sta alla base della nuova meccanica, segue in particolare che, se si calcola qual è il valor medio, $\langle A\rangle_x$, in un certo stato, rappresentato dal vettore normalizzato x, per l'osservabile rappresentata dall'operatore autoaggiunto A, si ottiene il prodotto interno $(x|Ax)$ (si ricordi che è questo un numero reale): esso, nell'ipotesi di semplicità adottata piú sopra, è pari evidentemente alla media $\sum_i a_i|w_i|^2$; lo scarto quadratico medio da tale valor medio è a sua volta definito dalla formula $\Delta_x A := (x|(A-\langle A\rangle_x)^2x)^{1/2}$; esso, per come è definito in base alle interpretazioni già ricordate, rappresenta l'incertezza con cui è noto il valore di una certa osservabile in un determinato stato; il caso piú «fortunato», come già si è visto, si ha quando il sistema si trova in un autostato dell'osservabile che si vuol misurare: si ha allora che il valor medio è esattamente uguale all'autovalore relativo all'autostato in questione e che l'incertezza $\Delta_x A$ è nulla: non vi è alcuna indeterminazione. Se ora si considerano due distinte osservabili rappresentate dagli operatori A e B è un puro esercizio di calcolo mostrare che vale sempre la disuguaglianza: $\Delta_x A \cdot \Delta_x B \geq (1/2)|(x|[A,B]x)|$; essa, alla luce dell'interpretazione che si è

descritta, è assai significativa: nel caso che i due operatori A e B commutino (osservabili compatibili), il secondo membro è zero e la disuguaglianza diventa banale poiché certamente il primo membro è un numero positivo o nullo; nel caso in cui i due operatori non commutino, il secondo membro, per un determinato stato, o per tutti gli stati possibili, può essere diverso da zero e ciò allora pone un *limite inferiore* al valore del primo membro; il prodotto delle due indeterminazioni non può essere inferiore, per quel determinato stato, a un certo valore; ciò implica, per quanto visto, che esse non possono essere contemporaneamente misurate con precisione arbitraria; assai semplice il caso della posizione e del momento lineare già menzionato: s'è visto che, per ogni componente, indicata dall'indice j, si ha $[Q_j,P_j]=i\hbar$; ciò evidentemente implica, per ogni stato x su cui ciò abbia senso, $\Delta_x Q_j \Delta_x P_j \geq \hbar/2 = h/4\pi$. Sono queste le celebrate relazioni d'indeterminazione di Heisenberg (*Unbestimmtheitsrelationen* o *Ungenauigkeitsrelationen*) delle quali Heisenberg [1930] è un'assai acuta e lucida illustrazione, soprattutto dal punto di vista dell'analisi delle esperienze in linea di principio effettuabili e delle limitazioni inevitabilmente ad esse inerenti. «Le relazioni [d'indeterminazione]... forniscono i limiti entro i quali possono essere impiegati i concetti della teoria corpuscolare. La ricerca di un piú esatto significato delle parole posizione e velocità di là dai limiti segnati dalle equazioni... è vana, come l'impiego di parole il cui senso non è stato definito» scrive Heisenberg [1930], e aggiunge in nota: «Occorre perciò concludere che il linguaggio umano permette in generale di formulare proposizioni, dalle quali non si potrebbe trarre alcuna conseguenza e che sono quindi completamente prive di significato; a esse però corrisponde nella nostra mente una specie di rappresentazione intuitiva. Cosí, per esempio, non può seguire alcuna conseguenza dall'asserzione che accanto al nostro mondo ne esista un secondo, col quale però non sia possibile avere mai, per ragioni di principio, alcuna relazione; ciò nonostante, c'è nella nostra fantasia una specie d'immagine corrispondente a quest'asserzione. Com'è evidente, una tale proposizione può essere con egual diritto sostenuta e contraddetta» (trad. it. pp. 26-27).

6.3. Dinamica quantistica.

Non s'è parlato finora di «leggi del moto»; la cinematica descrive solo il linguaggio formale che viene utilizzato per descrivere gli oggetti della teoria: i sistemi fisici.

Occorrerà osservare anzitutto che l'evoluzione nel tempo dello stato di un sistema fisico obbedisce a una equazione del moto fintantoché il sistema non subisce imprevedibili perturbazioni dall'esterno: tra queste tipicamente le osservazioni o *misure*; un sistema fisico obbedisce cioè a una fissata legge di movimento solo in periodi durante i quali non vengono effettuate su di esso osservazioni di sorta, e ciò coerentemente con quanto già si è visto a proposito della non completa controllabilità di un'osservazione: un tale processo, si è visto, altera il sistema in una misura che non è in linea di principio esattamente prevedibile. Ciò non esclude, anzi è in accordo con il fatto che, come già si è accennato, la

teoria permette di fare asserzioni probabilistiche (eventualmente anche con probabilità unitaria) sui risultati ottenibili nell'esecuzione di una misura di una determinata osservabile sul sistema che si trova in un determinato stato; è anzi naturalmente in virtú di questa circostanza che la teoria è sperimentalmente corroborabile. L'elemento cruciale che tuttavia distingue l'evoluzione del sistema imperturbato da osservazioni, dalla sua evoluzione «durante» un'osservazione, è precisamente questo, che mentre nella prima situazione esso evolve secondo una legge perfettamente determinata e deterministica, esattamente come nella meccanica classica, nella seconda situazione il sistema esibisce un comportamento governato soltanto da leggi probabilistiche. Per dare un'idea di quale sia la legge del moto che in generale governa l'evoluzione di un sistema quantistico occorre ricordare – per rimanere aderenti allo schema formale che qui si è accennato – che se A è un generico operatore autoaggiunto in uno spazio di Hilbert (e tali sono – si ricorderà – gli operatori che rappresentano osservabili) è possibile costruire la famiglia di operatori $U(t)$, per ogni valore reale del parametro t, indicata usualmente con la notazione $U(t)=\exp(iAt)$, in modo tale che, nel caso di spazi finito-dimensionali, coincida con l'esponenziazione della matrice (iAt): s'intende con la scrittura $\exp(iAt)$ la somma, sempre convergente, della serie $\sum_{n=0}^{\infty}((it)^n/n!)\,A^n$. Anche nel caso di spazi di Hilbert non finito-dimensionali, nel caso in cui l'operatore A, oltre ad essere autoaggiunto, sia anche limitato, la serie soprascritta è una buona definizione dell'operatore $U(t)$; definizioni piú generali, che passano sostanzialmente attraverso il teorema spettrale per gli operatori autoaggiunti, sono necessarie per operatori non limitati; si ricorderà qui di passaggio che tali situazioni non sono affatto «patologiche» o altamente improbabili nella teoria quantistica, e che ad esempio gli operatori posizione e momento lineare in $L^2(\mathbf{R})$ sono per l'appunto autoaggiunti ma non limitati. Anche in tali situazioni tuttavia, sviluppi in serie come quello scritto piú sopra possono conservare senso per particolari famiglie di vettori appartenenti al dominio dell'operatore A. Quel che qui interessa è comunque che la famiglia di operatori $U(t)$ possiede alcune notevoli proprietà che la rendono particolarmente adatta a realizzare l'evoluzione temporale di uno stato in meccanica quantistica. La prima è che, per ogni reale t, l'operatore $U(t)$ è unitario nello spazio di Hilbert, il che significa che si tratta di un automorfismo dello spazio (è cioè una biiezione che conserva la struttura: linearità e prodotto interno); si ha quindi per ogni valore di t e per ogni coppia di stati x, y dello spazio di Hilbert $\mathfrak{H}$, $(U(t)x \mid U(t)y)=(x|y)$, oltre che $U(t)(x+\alpha y)=U(t)x+\alpha\, U(t)y$; si ha inoltre per la famiglia $U(t)$: $U(O)=1$ e $U(t_1+t_2)=U(t_1)U(t_2)$ per ogni coppia di valori t_1 e t_2. Vale anche una notevole proprietà di continuità che si ometterà qui di precisare. Queste considerazioni formali, com'è ovvio, non hanno in sé alcun contenuto fisico, nel senso che non implicano, da sole, alcuna particolare legge del moto: quello che emerge è soltanto il fatto che una tale famiglia di operatori $U(t)$, detta gruppo continuo a un parametro di operatori unitari, è formalmente adatta a rappresentare l'evoluzione di un sistema fisico: se x_0 è lo stato di un tale sistema all'istante $t=0$, assunto convenzionalmente come iniziale, il vettore $x_t := U(t)x_0$ è un buon candi-

dato a rappresentare lo stato del sistema all'istante di tempo t. Occorre tuttavia, in maniera cruciale, individuare quale gruppo a un parametro $U(t)$ è quello che va inserito nella teoria; a tal fine vari ragionamenti sono stati introdotti per rendere plausibile la scelta che viene effettivamente fatta [cfr. ad esempio Dirac 1930, trad. it. p. 152]; nessuno di essi è realmente convincente se si guarda alla meccanica quantistica come teoria autonoma rispetto alle teorie precedenti, e in particolare alla meccanica newtoniana; si tratta in verità di un postulato, in favore del quale possono venir trovate varie argomentazioni di plausibilità riferite alle analogie con la meccanica classica. Un modo possibile (debole) per enunciarlo è il seguente: esiste un operatore autoaggiunto H, che rappresenta l'osservabile energia del sistema considerato, tale che il gruppo a un parametro $\exp(-(i/\hbar)Ht)$ rappresenta l'evoluzione dinamica del sistema; o tale che, se x_0 rappresenta lo stato del sistema all'istante $t=0$, lo stato del sistema all'istante t è rappresentato dal vettore $x_t=\exp(-(i/\hbar)Ht)x_0$. Una forma piú forte di questo postulato deve dare delle indicazioni sulla natura di questo fondamentale operatore energia: deve dare prescrizioni per costruirlo. I soli criteri che si possiedono per costruire l'operatore energia di un dato sistema, cruciale per individuare la sua evoluzione dinamica, sono quelli contenuti in una sorta di estensione del già citato principio di corrispondenza; i criteri di questa estensione sono già apparsi quando si è parlato degli operatori che rappresentano le osservabili posizione e momento lineare per la «particella in una dimensione»; si tratta infatti dei criteri che devono permettere, in modo del tutto generale, di pervenire all'espressione esplicita dell'operatore autoaggiunto che rappresenta una determinata osservabile; si è visto in quel caso che all'osservabile classica posizione, che si può indicare con la lettera x, si associava l'operatore autoaggiunto (s'intende qui nello spazio di Hilbert $L^2(\mathbf{R})$) Q, e all'osservabile classica momento lineare si associava l'operatore autoaggiunto P; una prescrizione notevolmente generale è allora quella di associare all'osservabile classica $r=f(q,p)$ l'osservabile quantistica, ovvero l'operatore autoaggiunto $R=f(Q,P)$, essendo la funzione f che compare in questa seconda definizione esattamente la stessa, dal punto di vista della sua forma, di quella che compare nella espressione che fornisce la r. Questa prescrizione è sufficientemente generale per lo scopo di fornire un equivalente quantistico di ogni osservabile classica, poiché si potrebbe dimostrare – pur di aggiungere opportune condizioni formali – che ogni osservabile classica è funzione della posizione e del momento lineare; non si tratta invece, in alcuni casi, di una prescrizione sufficientemente univoca, a causa del fatto che le funzioni di operatori sono oggetti piú delicati da trattare che non le funzioni di variabili reali; a parte infatti le difficoltà formali di definire una generica funzione di un solo operatore autoaggiunto, difficoltà superabili sempre mediante l'uso del teorema spettrale, quando si ha a che fare con una coppia di operatori che non commutano – quali sono Q e P – appaiono inevitabili ambiguità: tipicamente, se ci si chiede qual è l'operatore che corrisponde all'osservabile classica qp s'incontrano subito difficoltà in quanto QP non è autoaggiunto, e inoltre non si vede perché non scegliere PQ, visto che qp può anche essere scritto pq; d'altra parte PQ è ben diverso da QP (essi non commutano) e nessuno dei due è

autoaggiunto; si potrebbe scegliere invece $(1/2)(QP+PQ)$ che ha l'importante proprietà di autoaggiuntezza, ma neppure questa scelta è scevra di difficoltà. Si può obiettare che in realtà qp è un'osservabile solo in un senso molto indiretto, in quanto la sua misura dipende dalle misure di p e di q, d'altra parte sarebbe auspicabile possedere qualche regola generale per costruire le osservabili quantistiche corrispondenti a una vasta classe di osservabili classiche.

Questo problema non possiede alcuna soluzione soddisfacente; possiede in verità una soluzione pragmatica e una soluzione astratta. La prima soluzione – assai utile ai fini pratici di calcolo e di confronto con l'esperienza – consiste nell'esaminare, per ogni singolo sistema fisico con cui si tratta, quali sono le osservabili cui si è interessati – in primo luogo l'energia – e di trovare per esse un'espressione priva di ambiguità; ciò è il piú delle volte possibile in modo soddisfacente e permette quindi una impostazione precisa del problema quantistico. Tipicamente l'energia di un sistema conservativo consta usualmente della somma di due parti, l'energia cinetica e l'energia potenziale. Nella grande maggioranza dei casi, la prima è funzione soltanto del momento lineare dei corpi che costituiscono il sistema, mentre la seconda (il sistema è conservativo) è funzione soltanto della configurazione, e cioè delle posizioni rispettive di questi corpi; in una tale situazione ogni ambiguità è assente poiché la funzione $f(p,q)$ è della forma $f_1(p)+f_2(q)$, cosí che il corrispondente operatore quantistico avrà necessariamente la forma $f_1(P)+f_2(Q)$, nella quale non interviene mai un prodotto di operatori non commutanti; rientrano in questa situazione tutti i sistemi atomici semplici, per i quali possono cosí venir fatte previsioni quantitative molto precise, sempre in ottimo accordo con l'esperienza.

Vi sono certamente, in questo modo di fare, aspetti poco soddisfacenti dal punto di vista concettuale: ai fini di una sistemazione rigorosa della teoria si richiede di conoscere esattamente la delimitazione dell'insieme delle osservabili della teoria di un dato sistema fisico, e parallelamente si vorrebbe precisare con chiarezza la natura dei legami della nuova teoria con la meccanica classica; la soluzione astratta cui si accennava consiste nel conferire piena indipendenza, almeno nelle intenzioni, alla nuova teoria, studiando autonomamente le proprietà dell'insieme delle osservabili di un dato sistema, strutturato in algebra non commutativa con opportune proprietà. Non ci si addentrerà qui ovviamente in un tale terreno che richiederebbe ben altro spazio e ben altri strumenti formali.

Sarà opportuno d'altra parte insistere sul fatto che una volta noto, per un assegnato sistema fisico, l'operatore che rappresenta la sua energia totale, operatore usualmente denominato hamiltoniana del sistema, la sua evoluzione temporale è in linea di principio perfettamente determinata. Una forma praticamente equivalente nella quale può essere messa l'equazione che caratterizza tale evoluzione è la seguente: $i\hbar(d/dt)\,x_t=Hx_t$; in una forma simile a questa essa è stata per la prima volta scritta da Schrödinger ed è nota come equazione di Schrödinger dipendente dal tempo per le ragioni che diventeranno chiare tra breve; si tratta della forma differenziale dell'equazione scritta in forma intera piú sopra; anche da questa forma traspare il fatto che per la conoscenza dello stato del sistema all'istante t, x_t, è necessario conoscere, oltre ovviamente alla hamiltoniana

H, lo stato all'istante 0, x_0, che costituisce la condizione iniziale necessaria e sufficiente a risolvere una tale equazione differenziale.

Si osserverà che questa forma della legge dell'evoluzione dinamica di un sistema rende evidente il teorema di conservazione dell'energia, che conserva in meccanica quantistica tutta la sua validità, pur di venire opportunamente formulato: se un sistema fisico si trova in un certo istante (ad esempio $t=0$) in uno stato nel quale l'energia ha un valore definito, si trova cioè in un autostato dell'energia corrispondente all'autovalore E_0 per questa osservabile, il sistema permane, in assenza ovviamente di perturbazioni esterne, in tale situazione, cioè in uno stato in cui il valore dell'energia è ancora ben definito ed è uguale a E_0. È questa una conseguenza formale del fatto che $\exp((-i/\hbar)Ht)$, per ogni valore di t, è un operatore che commuta con H, in quanto è una sua funzione. Ciò suggerisce un risultato assai piú generale: se R è un qualsiasi operatore autoaggiunto *che commuta con* H, e che rappresenta una certa osservabile fisica del sistema, che viene supposto per semplicità possedere autovalori ed autovettori, allora anzitutto si può mostrare che si possono trovare autovettori simultanei di R e di H, diciamo relativi agli autovalori r_0 ed E_0 rispettivamente, ed inoltre si vede facilmente che se il sistema si trova a un certo istante in un tale autostato, in ogni istante successivo una misura, non solo dell'energia ma anche dell'osservabile rappresentata da R, fornirà esattamente i valori rispettivamente E_0 e r_0. Ovvero: le osservabili rappresentate da operatori che commutano con la hamiltoniana sono costanti del moto. Gli stati che hanno la proprietà detta rispetto alla hamiltoniana sono detti stati stazionari: essi rivestono un'importanza fondamentale nella teoria quantistica e nelle sue applicazioni, in quanto rappresentano gli stati stabili di un sistema; se un sistema si trova in uno stato stazionario, in esso permane fino a che non intervenga una perturbazione esterna (eventualmente una misura di una osservabile che non commuta con la hamiltoniana). La ricerca degli stati stazionari per un sistema fisico è uno dei problemi fondamentali della meccanica quantistica, in quanto essa fornisce l'insieme degli stati nei quali il sistema può stabilmente permanere; tale ricerca equivale, per quanto si è visto, a risolvere il problema agli autovalori per la hamiltoniana $Hx_r = E_r x_r$: problema matematicamente ben noto già prima dell'avvento della meccanica quantistica (cui anzi, si ricorderà, si riferiva Schrödinger nella sua ricerca) e la cui soluzione dipende fortemente dal tipo di operatore autoaggiunto che rappresenta l'energia del sistema. Una tale analisi non verrà qui proseguita (per qualche caso particolare ci si riferirà all'articolo «Moto» di questa stessa *Enciclopedia*). Si rileverà soltanto che, anche in casi molto semplici, il problema agli autovalori può non essere risolubile in senso stretto: può cioè accadere che una hamiltoniana anche di sistemi molto semplici *non ammetta alcun autovalore*: non esiste alcuno stato in cui l'energia è perfettamente determinata. Di questo tipo è certamente il sistema fisico piú semplice di tutti: la particella libera, sia in una sia in tre dimensioni; la sua hamiltoniana è data (si ricordi l'espressione classica dell'energia cinetica) dall'operatore $H=(1/2m)P^2$, dove m è la massa della particella e P l'operatore che corrisponde al suo momento lineare, diciamo in $L^2(\mathbf{R}^3)$; un tale operatore ha spettro puramente continuo ed esso è costituito dal semiasse reale positivo. Ciò significa che, da un lato, non vi sono alcun numero reale E e alcun elemento f di $L^2(\mathbf{R}^3)$ tali che si abbia $Hf=Ef$; significa d'altra parte che per ogni reale positivo E si può trovare un f in $L^2(\mathbf{R}^3)$ tale che Hf sia «vicino» a Ef quanto si vuole (preso cioè un ε positivo arbitrariamente piccolo, si può trovare un f (si tratta delle funzioni costanti in intervalli molto piccoli e nulle altrove cui s'è già fatto cenno sopra a p. 496) di norma unitaria tale che $\|Hf-Ef\|<\varepsilon$; in altre parole esistono stati che approssimano bene quanto si vuole un autostato della hamiltoniana relativo a un qualsiasi valore reale positivo).

Esistono altre situazioni in cui vi sono soltanto autovalori in senso proprio per la hamiltoniana (oscillatore armonico) ed altre ancora in cui vi sono sia autovalori sia spettro continuo.

Già a questo punto, come si vede, molta strada s'è fatta rispetto alle prime rozze ipotesi che «l'energia esista solo in quanti discreti» o che «possa solo venir scambiata in quanti discreti»; la rottura del 1900 ha già cambiato volto e le sue implicazioni sono andate straordinariamente piú in là di quanto mai s'immaginasse nei primi anni del secolo. La teoria che ha preso il via nel 1925 dal lavoro di Born, Heisenberg, Jordan e Schrödinger ha conosciuto negli anni successivi uno sviluppo probabilmente senza precedenti nella storia della fisica: la si è applicata con successo ad innumeri sistemi atomici per trarne previsioni e risultati da confrontare con gli innumeri esperimenti eseguiti e per proporre nuovi esperimenti; si sono investigate, e si stanno ancora investigando, le sue basi teoriche per capire su quali postulati potesse veramente venire fondata, per capire se e come si possa resecare il cordone ombelicale che sembra ineludibilmente legarla alla meccanica classica (ovvero per studiare tale legame), per esaminarne gli aspetti piú propriamente epistemologici che la caratterizzano in maniera assai peculiare rispetto alle teorie note prima della sua nascita, per capire infine se e come possano venire allargati i suoi orizzonti, superati i suoi piú o meno apparenti paradossi, estese le sue applicazioni ad altri settori dell'esperienza.

Mentre è qui necessario arrestarsi alle soglie di questa ormai sterminata problematica, v'è un'ultima domanda cui si vorrebbe tentare di fornire risposta, se mai una tale risposta esiste. Ed è la domanda che chiede d'individuare le caratteristiche essenziali della nuova teoria, d'illuminare quel nocciolo duro della meccanica quantistica che dovrebbe distinguerla da ogni altra teoria e che dovrebbe resistere inalterato e protetto attraverso tutta la storia della teoria e rimanere sottostante a tutte le interpretazioni e a tutte le versioni che della meccanica quantistica sono ormai state date.

7. *Il nocciolo duro della meccanica quantistica.*

Quel che finora s'è detto della meccanica quantistica non permette certamente una disamina articolata di tutti gli aspetti importanti della teoria, molti dei quali non sono neppure stati menzionati, né permette un'analisi abbastanza fine delle contraddizioni insite in alcune interpretazioni della teoria e della loro possibile eliminazione [cfr. ad esempio Bohr 1949; Wigner 1973].

Va prima d'ogni altra cosa sottolineato che il nocciolo duro della meccanica quantistica non consisterà d'una singola affermazione, o equazione, o postulato, ma che, se esso è in qualche modo individuabile, consisterà di un insieme interconnesso di vari di questi ingredienti. La già citata, e paradossale, definizione feyerabendiana, riferita per la verità alla sola interpretazione di Copenhagen, esprime peraltro l'impressione che può ricavarsi da un atteggiamento troppo complessivo nei riguardi della meccanica quantistica, cosí come di fatto si è venuta sviluppando nell'ultimo mezzo secolo.

Si prescinderà qui anche dall'esaminare i cosiddetti approcci assiomatici alla meccanica quantistica che tuttora, nella loro notevole varietà e differenziazione, non sembrano ancora maturi per cogliere con sufficiente completezza il nocciolo della teoria; nocciolo che del resto va cercato *prima* di tale tipo di approcci, che, al meglio, si propongono di coglierlo e di formalizzarlo completamente.

Si comincerà con il ricordare come una delle esigenze che piú fortemente portò alla nascita della nuova meccanica sia stata quella di avere una nuova cinematica, cioè una nuova descrizione del moto di un sistema atomico, nella misura in cui il modo classico di descrivere un sistema fisico non sembrava adeguato a questo scopo. A sua volta, questa esigenza scaturiva dalla convinzione che la teoria doveva parlare soltanto delle grandezze osservabili dei sistemi in gioco e che nulla doveva essere detto, in quanto privi di significato, di quei concetti che non potevano, per ragioni di principio, essere oggetto di esperienza. Questo principio ebbe in verità un enorme valore euristico durante tutto lo sviluppo della nuova teoria e potrebbe certamente venire annoverato come uno degli elementi essenziali dello *hard core* della meccanica quantistica; se non fosse che, come facilmente si capisce, esso è difficilmente precisabile, quando si viene al concreto della teoria e delle esperienze che con essa si pensano connesse. Vi sono esperimenti concettuali, in qualche caso assai autorevoli [si veda per tutti Einstein, Podolski e Rosen 1935], che hanno fatto apparire varie aporie alla base della teoria; o anche, si potrebbe dire, ne hanno messo in luce caratteristiche affatto peculiari, nel caso citato la inevitabile non-completezza, e che sono peraltro stati fonti di lungo dibattito e di non risolto disaccordo tra i piú autorevoli protagonisti. Un punto fondamentale a questo riguardo sembra essere il fatto che la decisione su quali sono gli esperimenti in linea di principio ammessi e quelli che non lo sono è inevitabilmente dipendente dalla teoria stessa. Si badi che ciò non sminuisce neppur di poco l'importanza del principio quale ispiratore di modelli di sviluppo della teoria, ma semplicemente costituisce un primo supporto della tesi, che si vorrebbe qui introdurre, sulla particolare e peculiare difficoltà connessa con l'individuazione di uno *hard core* per la meccanica quantistica.

Non molta fortuna avrebbe il tentativo di far coincidere tale nucleo con il formalismo matematico che sorregge tutta la teoria: lasciando per un attimo da parte il fatto, peraltro cruciale, che tale formalismo non determina l'interpretazione che lo rende connesso con la teoria, basta ricordare quanto non unico sia tale formalismo: l'originale formulazione hilbertiana di Neumann, che anche in queste pagine è stata usata, non è certo rimasta tale; essa è stata certamente perfezionata ed arricchita per vari scopi e applicazioni, soprattutto però ad essa sono

state sostituite, e spesso con vantaggio, altre formalizzazioni, fondate ad esempio sulla teoria delle algebre (anche qui con possibili diversificazioni), o di altre strutture algebriche, e non v'è modo, a tutt'oggi, di privilegiare nettamente una di queste formulazioni rispetto alle altre, senza perdere aspetti spesso essenziali che alla teoria vengono da queste varie fonti.

Se poi si prendono in considerazione le cosiddette «interpretazioni», ovvero quella parte essenziale della teoria che dovrebbe connettere le formule che ne costituiscono la formalizzazione con le proposizioni osservative che alla teoria si pensano particolarmente connesse, il pluralismo aumenta. Probabilmente la prima e piú importante disparità si manifestò a proposito di quale dovesse essere l'*oggetto* tipico della meccanica quantistica: mentre, anche qui la terminologia e le designazioni sono un po' riassuntive, l'interpretazione piú «ortodossa» tendeva a insistere affinché la nuova teoria si applicasse anche a un singolo sistema atomico, facendo su di esso predizioni, naturalmente entro i limiti intrinseci alla teoria, del tutto valide, si affacciò presto un'interpretazione secondo la quale essa poteva venire applicata soltanto ad insiemi statistici di sistemi, ovvero a numeri molto grandi di sistemi identici a quello in esame e tutti «preparati» nello stesso modo. Distinzione certo molto importante dal punto di vista della costituzione della meccanica quantistica in teoria complessiva dei fenomeni microscopici, meno forse dal punto di vista della prassi di chi ne applicava le norme, o le prescrizioni algoritmiche per calcolare i livelli energetici di qualche sistema atomico, o per calcolare le sezioni d'urto di qualche reazione tra particelle. In tali prassi la teoria viene di fatto applicata il piú delle volte ad insiemi molto numerosi di sistemi identici, anche se, come mostra Espagnat [1976, trad. it. p. 133], è perfettamente coerente pensare che almeno certi tipi di stati fisici di un singolo sistema vengano descritti da un opportuno vettore di uno spazio di Hilbert.

Forse una delle caratteristiche piú controverse, e che sempre sono apparse piú ineludibili della nuova teoria, è stata la sua non-completezza. S'intende qui per il momento con tale espressione l'intrinsecamente minore ricchezza di informazioni che la teoria prevede di trattare per un sistema fisico. È noto che in uno stato, che può pensarsi individuato da valori fissati di un certo numero di osservabili tra loro commutanti, non è possibile affermare un valore preciso per una qualsiasi altra osservabile che non commuti con quelle che lo individuano. Se un oscillatore armonico si trova in un preciso autostato della sua energia, quindi in uno stato stazionario, non è possibile dire quale posizione nello spazio esso occupa, né quale momento lineare possiede; è possibile soltanto, a proposito di queste osservabili, fare affermazioni probabilistiche, ovvero dire quali sono i loro *valori medi*. Si tratta certamente di una rinunzia assai pesante e fu probabilmente alla base delle motivazioni che spinsero Einstein a non accettare mai la meccanica quantistica come una teoria soddisfacente per la descrizione fisica dei fenomeni. Il famoso passo della sua lettera del dicembre 1926 a Born [Einstein e Born 1969] «La meccanica quantistica è degna di ogni rispetto, ma una voce interiore mi dice che non è ancora la soluzione giusta. È una teoria che ci dice molte cose, ma non ci fa penetrare piú a fondo il segreto del gran Vecchio. In ogni caso, sono convinto che *questi* non gioca a dadi col mondo» (trad. it.

pp. 108-9), non fa che esprimere in maniera pittoresca un atteggiamento già manifestato nell'aprile del 1924: «Le idee di Bohr sulla radiazione m'interessano molto, ma non vorrei lasciarmi indurre ad abbandonare la causalità rigorosa senza aver prima lottato in modo assai diverso da come s'è fatto finora. L'idea che un elettrone esposto a una radiazione possa scegliere *liberamente* l'istante e la direzione in cui spiccare il salto è per me intollerabile. Se cosí fosse, preferirei fare il ciabattino, o magari il biscazziere, anziché il fisico. Vero è che i miei tentativi di dare una forma tangibile ai quanti sono regolarmente abortiti, ma sono ben lontano dal rinunziare a questa speranza. E anche se la cosa non riuscirà, mi rimarrà sempre la consolazione di aver fallito da solo» [*ibid.*, p. 98]. Assai significativo per comprendere le profonde contraddizioni filosofiche che sottostavano a queste affermazioni è leggere il commento di Born stesso a questa lettera a lui diretta: «Troviamo qui una professione di fede fisico-filosofica: il rifiuto di considerare le leggi statistiche come fondamenti ultimi della fisica... Einstein era fermamente convinto che la fisica ci fa accedere alla conoscenza di un mondo esterno oggettivamente esistente; lo studio dei fenomeni quantici dell'atomo ha invece persuaso a poco a poco me e molti altri fisici che le cose non stanno cosí. In ogni istante abbiamo solo una conoscenza rudimentale e approssimata del mondo oggettivo, e da questa conoscenza possiamo trarre – secondo certe regole, che sono le leggi di probabilità della meccanica quantistica – certe conclusioni sugli stati ancora incogniti (ad esempio su quelli futuri)» [*ibid.*, p. 99]. È dunque questa intrinseca impotenza a descrivere «completamente» un sistema fisico una caratteristica determinante della teoria, tale anzi da aver suscitato in alcuni dei protagonisti di primo piano aspri rifiuti; lo stesso Schrödinger non accettò mai la versione di Göttingen-Copenhagen della teoria, pur avendo egli fornito, con il suo formalismo ondulatorio, un impulso essenziale allo sviluppo e alla diffusione della teoria; si ricordi del resto che l'interpretazione probabilistica della famosa ψ, la «funzione d'onda» di Schrödinger, fu data da Born, un fisico che, pur appartenendo alla generazione di Einstein (ne era solo di tre anni piú giovane), era caratterizzato da una grande maggior spregiudicatezza nei confronti del modo di vedere classico. Emblematicamente, l'idea cui forse Schrödinger era piú attaccato era quella di *continuità* in fisica! Egli non si risolse mai ad accettare i «salti quantici» di Bohr, né l'interpretazione di Born della sua teoria. Ancora nel 1950 si legge in una lettera di Einstein a Schrödinger: «Tu sottolinei del tutto giustamente che la descrizione completa non può venire fondata sul concetto di accelerazione e tanto meno – cosí a me sembra – sul concetto di particella. Tra gli strumenti a nostra disposizione non rimane che il concetto di campo; ma sa il diavolo se anche questo resisterà. Io penso che valga la pena di restar fedeli a tale concetto, che è come dire al continuo, finché almeno non si possieda una qualche solida motivazione contraria» [Schrödinger e altri 1926-50, pp. 36-37]. Né Einstein né Schrödinger nel 1950 vedevano alcuna solida ragione per una rottura che era cominciata mezzo secolo prima.

Tuttavia, anche la non-completezza della teoria costituisce un aspetto tutt'altro che precisato e assoggettabile a una formulazione definitiva. Si osserverà anzitutto che una simile caratteristica assume rilevanza diversa a seconda che si

pensi la meccanica quantistica come una descrizione di un sistema o di un numero molto grande di sistemi; ed in secondo luogo si ricorderà qui che anche a tale non-completezza s'è cercato di porre rimedio mediante il formalismo delle cosiddette «variabili nascoste». Le teorie a variabili nascoste, che qui non si descriveranno, prendono la loro origine dall'assunzione che la non-completezza della meccanica quantistica non sia inevitabile: si suppone cioè che la meccanica quantistica non sia *tutta la verità*, ma parte di essa: esisterebbe un modo d'introdurre un maggior numero di parametri nella teoria – le variabili nascoste – che trattino anche le grandezze osservabili di cui la meccanica quantistica dichiara di non poter prevedere i valori, e che rendono dunque la teoria completa nel senso che era familiare dalle teorie classiche; la meccanica quantistica non sarebbe che la punta di un iceberg, motivato e sorretto da un apparato sottostante perfettamente deterministico e causale. In questo modo ogni sistema fisico torna ad avere proprietà definite sue proprie, ed i valori medi calcolati tenendo conto delle variabili nascoste riproducono esattamente quelli previsti dalla meccanica quantistica. Una formulazione a variabili nascoste fu costruita in modo coerente da Bohm [1952] e portata avanti e discussa da Bell [1966]. Certamente una tale teoria conosce non poche difficoltà – per esempio una forte non-località, vi sono correlazioni tra eventi spazialmente assai lontani – ma sembra del tutto compatibile con gli aspetti piú strettamente pertinenti alla meccanica quantistica standard.

Quanto s'è accennato nelle ultime pagine non deve far pensare naturalmente che la meccanica quantistica non esiste: essa esiste massicciamente e comprende un insieme di metodi, di concetti, di procedure di calcolo e di risultati numerici che le hanno permesso d'invadere negli ultimi cinquant'anni praticamente tutta la fisica e notevoli parti della chimica e della biologia. Quello che sembra assai piú dubbio è che si possa individuare un nocciolo duro della teoria non soggetto a trasformazioni e non contraddetto da esperimenti. Se si guarda alla meccanica quantistica cosí come realmente si è sviluppata fino ad oggi e cosí come è stata praticata dai suoi ormai numerosissimi adepti viene piuttosto in mente la definizione kuhniana di *esemplare* [Kuhn 1973, p. 463] come di «concrete soluzioni di problemi, accettate dal gruppo come paradigmatiche, nel senso del tutto usuale della parola». Il livello assiomatico è ancora, come s'è detto, assai lontano dall'esser raggiunto, e piú in generale una sistemazione rigorosa della teoria che elimini tutte le contraddizioni e risponda a tutte le domande, con essa compatibili beninteso, non esiste. Lo stesso procedimento di quantizzazione, cosiddetto, cioè quel procedimento che dovrebbe permettere di passare dalla descrizione classica di un dato sistema fisico alla sua descrizione quantistica, è certamente noto in una serie assai numerosa di casi particolari – si tratta di un complesso assai ricco di esemplari – ma è ancora oggetto di ricerca nella sua generalità. La rottura drastica che costituí la molla di partenza della vecchia teoria quantistica, e quindi della meccanica quantistica, rottura dovuta, si noti, a Planck e ad Einstein, è andata soggetta a un tale rimescolamento di carte da essere, a cose fatte, irriconoscibile. [A.S.].

Bell, J. S.
1966 *On the problem of hidden variables in quantum mechanics*, in «Reviews of Modern Physics», XXXVIII, 3, pp. 447-52.

Bohm, D.
1952 *A suggested interpretation of the quantum theory in terms of "hidden" variables. I*, in «Physical Review», LXXXV, 2, pp. 166-79.

Bohr, N.
1918 *On the quantum theory of line-spectra*, in «Kongelige Danske Videnskabernes Selskabs Skrifter, Naturvidenskabelig og Mathematisk Afdeling», serie VIII, IV, pp. 1-36.
1922 *Über die Anwendung der Quantentheorie auf den Atombau*, I. *Die Grundpostulate der Quantentheorie*, in «Zeitschrift für Physik», XIII, pp. 117-65.
1949 *Discussion with Einstein on Epistemological Problems in Atomic Physics*, in P. A. Schilpp (a cura di), *Albert Einstein: Philosopher-Scientist*, The Library of Living Philosophers, Evanston Ill. (trad. it. Boringhieri, Torino 1958, pp. 147-90).
1961 *Die Entstehung der Quantenmechanik*, in F. Bopp (a cura di), *Werner Heisenberg und die Physik unserer Zeit*, Vieweg, Braunschweig, pp. IX-XII.

Bohr, N.; Kramers, H. A.; e Slater, J. C.
1924 *The quantum theory of radiation*, in «Philosophical Magazine», serie VI, XLVII, pp. 785-802.

Born, M.
[1940-63] *My Life. Recollections of a Nobel Laureate*, Taylor and Francis, London - Scribner's, New York 1978 (trad. it. Editori Riuniti, Roma 1980).

Born, M., e Jordan, P.
1925 *Zur Quantenmechanik*, in «Zeitschrift für Physik», XXXIV, pp. 858-88.
1929 *Elementare Quantenmechanik*, Springer, Berlin.

Born, M.; Heisenberg, W.; e Jordan, P.
1925 *Zur Quantenmechanik II*, in «Zeitschrift für Physik», XXXV, pp. 557-615.

Broglie, L. de
1924 *Recherches sur la théorie des quanta*, Masson, Paris.

Dijksterhuis, E. J.
1950 *De Mechanisering van het Wereldbeeld*, Meulenhoff, Amsterdam (trad. it. Feltrinelli, Milano 1971).

Dirac, P.-A.-M.
1926*a* *The fundamental equations of quantum mechanics*, in «Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences», CIX, pp. 642-53.
1926*b* *Quantum mechanics and a preliminary investigation of the hydrogen atom, ibid.*, CX, pp. 561-79.
1930 *The Principles of Quantum Mechanics*, Clarendon Press, Oxford 1958[4] (trad. it. Boringhieri, Torino 1959).

Einstein, A.
1905 *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesischtspunkt*, in «Annalen der Physik», XVII, pp. 132-48 (trad. it. in *La teoria dei quanti di luce*, Newton Compton, Roma 1972, pp. 45-72).
1917 *Quantentheorie der Strahlung*, in «Physikalische Zeitschrift», XVIII, pp. 121-28.
[1946] *Autobiographical Notes*, in P. A. Schilpp (a cura di), *Albert Einstein: Philosopher-Scientist*, The Library of Living Philosophers, Evanston Ill. 1949 (trad. it. Boringhieri, Torino 1958, pp. 3-49).

Einstein, A., e Born, M.
1969 *Briefwechsel 1916-1955*, Nymphenburger Verlagshandlung, München (trad. it. Einaudi, Torino 1973).

Einstein, A.; Podolski, B.; e Rosen, N.
1935 *Can quantum-mechanical description of physical reality be considered complete?*, in «Physical Review», XLVII, pp. 777-80.

Espagnat, B. d'
1976 *Conceptual Foundations of Quantum Mechanics*, Benjamin, Reading Mass. 1976[2] (trad. it. Bibliopolis, Napoli 1980).

Feyerabend, P. K.
1968 *On a recent critique of complementarity: part I*, in «Philosophy of Science», XXXV, pp. 309-31.

Hegel, G. W. F.
1830 *Encyklopädie der philosophischen Wissenschaften im Grundrisse*, Oswald, Heidelberg 1830[3] (trad. it. Laterza, Bari 1978[4]).

Heisenberg W.
1925 *Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen*, in «Zeitschrift für Physik», XXXIII, pp. 879-93.
1930 *Die physikalischen Prinzipien der Quantentheorie*, Hirzel, Leipzig (trad. it. Boringhieri, Torino 1979).
1964 *Quantum Theory and Its Interpretation*, in S. Rozental (a cura di), *Niels Bohr hans liv og virke fortalt af en kreds af venner og medarbejdere*, Schultz, København (trad. ingl. North-Holland, Amsterdam 1967, pp. 94-108).

Kelvin (W. Thomson, Lord)
[1884] *Baltimore Lectures on Molecular Dynamics and the Wave Theory of Light*, Clay, London 1904.

Klein, M. J.
1979 *Paradigm lost? A review symposium*, in «Isis», LXX, pp. 429-34.

Kuhn, Th. S.
[1973] *Second Thoughts on Paradigms*, in F. Suppe (a cura di), *The Structure of Scientific Theories*, University of Illinois Press, Urbana 1977[2], pp. 459-82.
1978 *Black-Body Theory and the Quantum Discontinuity 1894-1912*, Clarendon Press, Oxford.

Ladenburg, R.
1921 *Die Quantentheoretische Deutung der Zahl der Dispersionselektronen*, in «Zeitschrift für Physik», IV, pp. 451-68.

Landé, A.
1928 *Optik, Mechanik und Wellenmechanik*, in H. Geiger e K. Scheel (a cura di), *Handbuch der Physik*, XX. *Licht als Wellenbewegung*, Springer, Berlin, pp. 317-452.

Lewis, G. N.
1926 *The conservation of photons*, in «Nature», CXVIII, pp. 874-75.

Neumann, J. von
1932 *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin.

Pauli, W.
1933 *Die allgemeine Prinzipien der Quantenmechanik*, in H. Geiger e K. Scheel (a cura di), *Handbuch der Physik*, XXIV. *Quantentheorie*, Springer, Berlin 1933[2], pp. 83-272.

Planck, M.
1900*a* *Über irreversible Strahlungsvorgänge*, in «Annalen der Physik», serie IV, I, 1, pp. 69-122.
1900*b* *Entropie und Temperatur strahlender Wärme, ibid.*, 4, pp. 719-37.
1900*c* *Über eine Verbesserung der Wien'schen Spektralgleichung*, in «Verhandlungen der Deutschen Physikalischen Gesellschaft im Jahre 1900», II, 13, pp. 202-4.
1900*d* *Zur Theorie des Gesetzes der Energieverteilung im Normalspektrum, ibid.*, 17, pp. 237-45.
1943 *Zur Geschichte der Auffindung des physikalischen Wirkungsquantums*, in «Die Naturwissenschaften», XXXI, 14-15, pp. 153-59.

Poincaré, H.
1912 *La théorie des quanta*, in «Journal de Physique Théorique et Appliquée», serie V, II, pp. 5-34.

Rayleigh (J. W. Strutt, Lord)
1900 *Remarks upon the law of complete radiation*, in «Philosophical Magazine», XLIX, pp. 539-40.

Rosenfeld, L.
1970 *Men and ideas in the history of atomic theory*, in «Archive for History of Exact Sciences», VII, pp. 69-90.

Rubens, H., e Kurlbaum, F.
1900 *Über die Emission langwelliger Wärmestrahlen durch den schwarzen Körper bei verschiedenen Temperaturen*, in «Sitzungsberichte der Königlich-Preussischen Akademie der Wissenschaften zu Berlin», n. 41, pp. 929-41.

Schrödinger, E.
1926*a* *Quantisierung als Eigenwertproblem*, in «Annalen der Physik», LXXIX, pp. 361-76.
1926*b* *Über das Verhältnis der Heisenberg-Born-Jordanschen Quantenmechanik zu der meinen, ibid.*, LXXIX, pp. 734-57.

Schrödinger, E., e altri
[1926-50] *Briefe zur Wellenmechanik*, Springer, Wien 1963.

Slater, J. C.
1924 *Radiation and atom*, in «Nature», CXIII, p. 307.

Sommerfeld, A. J. W.
1929 *Atombau und Spektrallinien. Wellenmechanischer Ergänzungsband*, Vieweg, Braunschweig.

Waerden, B. L. van der
1967 (a cura di) *Sources of Quantum Mechanics*, North-Holland, Amsterdam.

Weyl, H.
[1917] *Das Kontinuum: kritische Untersuchungen über die Grundlagen der Analysis*, De Gruyter, Berlin-Leipzig 1932 (trad. it. Bibliopolis, Napoli 1977).

Wigner, E. P.
1973 *Epistemological Perspective on Quantum Theory*, in C. A. Hooker (a cura di), *Contemporary Research in the Foundations and Philosophy of Quantum Theory*, Reidel, Dordrecht, pp. 369-85.

Wittgenstein, L.
1922 *Tractatus logico-philosophicus*, Kegan Paul, Trench and Trubner, London (trad. it. Einaudi, Torino 1974).

Il punto di vista quantistico in **fisica**, nato dall'aggregazione di alcune specifiche soluzioni per problemi e anomalie del punto di vista «classico» (radiazione del corpo nero in termodinamica: cfr. **energia**, **entropia**, **reversibilità/irreversibilità**, effetto fotoelettrico: cfr. ad esempio **luce**), si è quindi articolato in un **paradigma** il cui doppio aspetto – «matriciale» (Born, Heisenberg, Jordan) e «ondulatorio» (Broglie, Schrödinger) – ha dato infine luogo a un complesso non necessariamente coerente di ipotesi, metodi e modelli (modificando sensibilmente la stessa idea di «teoria fisica»: cfr. del resto **ipotesi**, **metodo**, **modello**, **teoria/modello**; e, per il nesso teoria-esperienza, anche **empiria/esperienza**, **esperimento**) circa la struttura atomica (cfr. **atomo e molecola**) della **materia** (cfr. anche **moto**, **particella**).

Sotto il profilo piú generale, il punto di vista quantistico rappresenta una specializzazione – e vuol essere un superamento – dell'opposizione (cfr. **coppie filosofiche**, **opposizione/contraddizione**) **continuo/discreto**: l'estensione progressiva dei metodi quantistici ad altri settori della conoscenza della **natura** ha dunque sollevato non poche controversie di natura non solo scientifica (cfr. **ricerca**, **scienza**) ma anche filosofica, in cui si sono scontrate le «metafisiche influenti» delle varie «scuole» (cfr. **filosofia/filosofie**, e soprattutto **metafisica**), come in particolare mostra la ormai celebre disputa sul carattere **determinato/indeterminato** del **reale** (cfr. anche **caso/probabilità**). Ciò ha compor-

tato però notevoli sviluppi nelle stesse scienze della **vita**, mentre non ha mancato d'influenzare l'ambito delle **matematiche** specie nel contesto delle tecniche di **calcolo** impiegate (cfr. anche **applicazioni**, **funzioni** e, per altri aspetti, **strutture matematiche**): infine le formulazioni piú rigorose del punto di vista quantistico – movendo dal problema di un'assiomatica adeguata (cfr. **assioma/postulato**) – hanno retroagito sulla stessa **logica** (cfr. anche **deduzione/prova**, **formalizzazione**, **possibilità/necessità**, **ricorsività**), come le ricerche di «logica quantistica» attualmente stanno dimostrando con risultati sempre piú convincenti.

Relatività

Nel 1905 apparve sugli «Annalen der Physik» un articolo ad opera di Albert Einstein, all'epoca sconosciuto impiegato dell'ufficio brevetti di Berna, in cui venivano trattate a fondo alcune questioni riguardanti il moto dei corpi carichi elettricamente [Einstein 1905*a*]. Questo lavoro viene assunto quasi unanimemente come il punto di partenza della teoria della relatività, il cui dominio di applicazione si rivelò ben presto molto piú vasto ed universale di quello del problema da cui aveva tratto origine.

Infatti risulta chiaro che la teoria affonda le sue radici ed è sorretta dal rigoglioso sviluppo scientifico del XIX secolo; in particolare, la relatività non sarebbe concepibile senza la costruzione delle equazioni del campo elettromagnetico, ad opera di Maxwell, che di questo secolo furono il punto culminante.

Precedette il lavoro di Einstein una memoria di Lorentz [1904] in cui i problemi che sorgevano da certe anomalie di propagazione della luce (in particolare il risultato negativo della esperienza di Michelson-Morley) vengono correttamente impostati e in parte risolti. Come fa tuttavia notare Holton [1960], la teoria di Lorentz abbisogna di ben dodici distinte assunzioni, alcune delle quali artificiose ed improponibili in una formulazione piú moderna. Inoltre la teoria di Lorentz è corretta solo fino ai termini di ordine $(v/c)^2$, mentre quella di Einstein è rigorosamente valida a tutti gli ordini; mancano nella prima l'apertura e il senso di universalità che continuano ad affascinarci nella relatività.

Va tuttavia riconosciuto a Lorentz e a Poincaré il merito grandissimo di avere intuito l'importanza dei problemi che si presentavano nella teoria di Maxwell e di avere contribuito in modo sostanziale alla loro soluzione (trasformazioni di Lorentz). In particolare, il lavoro di Poincaré [1904], che precedette di poco quello di Einstein, riempie le lacune di quello di Lorentz impostando correttamente il problema dell'invarianza relativistica delle equazioni di Maxwell nel vuoto. Whittaker [1953] attribuisce il merito della scoperta a Poincaré e a Lorentz, relegando Einstein al ruolo di fortunato divulgatore della teoria: è un giudizio ingiusto e certamente influenzato da pregiudizi personali.

Negli anni che seguirono il famoso lavoro del 1905, la relatività cominciò ad essere gradualmente accettata e la sua portata ed estrema eleganza concettuale divennero sempre piú evidenti attraverso varie rielaborazioni, sia dello stesso Einstein, sia ad esempio di Minkowski [1908] che ebbe per primo una chiara idea del continuo spazio-tempo. Già nel 1905 fu anche chiaro che la teoria prevedeva l'equivalenza di massa ed energia, secondo la ormai celebre $E=mc^2$ [Einstein 1905*b*]. Lo sviluppo del calcolo tensoriale permise di snellire i calcoli, a scapito forse del loro contenuto intuitivo; la semplicità cosí raggiunta, tuttavia, ha permesso di scoprire profonde analogie formali che continuano ad essere operanti e a dare frutti nella fisica contemporanea [Woolf 1979].

Che dice dunque la relatività del 1905? Una risposta completa sarà data solo attraverso la trattazione, di necessità tecnica, contenuta nei paragrafi seguenti [per una conoscenza piú approfondita e per dettagli tecnici, si vedano Weinberg 1972; Misner, Thorne e Wheeler 1973]. Al momento si possono condensare per sommi capi i punti essenziali della teoria in quanto segue:

A) I fenomeni fisici si svolgono nell'ambito dello spazio $\mathbf{R}^3$ a tre dimensioni. Questo spazio contiene la materia e la configurazione di questa cambia nel tempo. Considerando quindi la storia completa dell'evoluzione dell'universo dal tempo $t=-\infty$ al tempo $t=+\infty$, è possibile considerarla come un insieme di eventi (caratterizzati dalla posizione in $\mathbf{R}^3$ e dal tempo t). Questi eventi formano un continuo $\mathfrak{M}$, lo spazio-tempo di Minkowski, che è l'ente fondamentale della teoria della relatività.

B) L'universo di Minkowski viene dotato di osservatori *ideali*, provvisti di orologi e di metri, e quindi capaci di misurare le coordinate di un evento qualsiasi in $\mathfrak{M}$. La relatività ristretta asserisce l'esistenza di una classe privilegiata di osservatori, detti *inerziali*, per i quali vale il principio d'inerzia. Tutti questi osservatori si muovono di moto relativo uniforme e un corpo sul quale non agiscono forze (corpo libero) si muove di moto uniforme rispetto ad essi. La relatività ristretta non è riconducibile ad assunzioni piú primitive, essa è un'idealizzazione che rende mirabilmente conto di una grande varietà di dati empirici e va accettata come tale. Esistono osservatori non inerziali, ad esempio una piattaforma rotante; la relatività ristretta non s'interessa ancora di essi, il compito è demandato a quella generale.

Conseguenza della relatività ristretta è che esistono infinite maniere diverse ed equivalenti di descrivere i fenomeni nello spazio-tempo. Certe variabili fisiche possono avere lo stesso valore per tutti gli osservatori (si diranno assolute o meglio scalari), altre, e sono la maggioranza, dipendono invece dall'osservatore. Queste ultime, il cui significato è relativo, si esprimono in genere mediante vettori e tensori. La formulazione di Einstein ha dimostrato che certe variabili, che si credevano assolute nel passato, sono in realtà relative: la differenza di tempo tra due eventi, la lunghezza di una sbarra, la forma di un oggetto sono, contrariamente all'intuizione, delle variabili relative.

Questa azione di rottura dell'intuizione tradizionale ha portato forse a porre troppa enfasi sul concetto di relativo, portando a formule superficiali quali «Tutto è relativo, nulla è assoluto». Si tratta sempre di una realtà unica, lo spazio-tempo, che viene vista sotto prospettive diverse da diversi osservatori. Tutte queste prospettive sono legate tra di loro in modo semplice, diretto e privo di ambiguità dalle trasformazioni di Lorentz e generalizzazioni, e sono perfettamente controllabili mediante il calcolo tensoriale.

C) Un'asserzione assoluta, nel senso discusso in B), riguarda la velocità della luce nel vuoto c, che risulta uguale (vale circa 300 000 km/sec) per tutti gli osservatori inerziali. È questo un risultato inaspettato, in quanto la velocità di un oggetto sembra un concetto relativo. Inseguendo la luce a 100 000 km/sec, essa dovrebbe fuggirci a 200 000 km/sec, la sua velocità dovrebbe risultare funzione dell'osservatore ed essere un concetto relativo. Queste considerazioni poggiano tuttavia sul teorema di addizione delle velocità, che risulta inconsistente con l'evidenza empirica e che viene sostituito nella relatività con una formula

modificata. Componendo con c una velocità qualsiasi si ritrova sempre c. Ma non è possibile assumere il carattere assoluto di c senza sconvolgere tutte le nozioni tradizionali dello spazio-tempo. Eventi simultanei per un osservatore non lo sono piú per un altro e moltissime proprietà assolute si scoprono relative.

Da questo punto di vista Einstein salvò il principio di relatività contenuto in B). Se c dipendesse dall'osservatore, si potrebbe operare una discriminazione tra osservatori inerziali, che distruggerebbe la completa simmetria ed equivalenza tra di essi ai fini della descrizione degli eventi. Queste simmetria ed equivalenza parvero molto piú importanti ad Einstein del teorema di addizione delle velocità: la relatività nacque da questa critica e non, come sostiene l'aneddotica tradizionale, dall'esperienza di Michelson-Morley.

Nel 1910 la struttura relativistica del campo elettromagnetico era stata chiarita in tutti i dettagli essenziali. Rimaneva invece oscuro il ruolo del campo gravitazionale. In quest'epoca Einstein era praticamente l'unico tra i grandi fisici che si rendesse conto dell'importanza del problema; negli anni seguenti, l'atomo di Bohr (si veda l'articolo «Atomo e molecola» in questa stessa *Enciclopedia*) e il dibattito sui quanti di luce [cfr. Schilpp 1949] contribuirono a distrarre i suoi colleghi dal campo gravitazionale.

La costruzione di una teoria relativistica della gravità impegnò Einstein dal 1908 fino al 1916, anno in cui finalmente raggiunse la meta dopo vari tentativi infruttuosi. Il risultato finale [Einstein 1916] va sotto il nome di «relatività generale»; questa nuova formulazione contiene come caso particolare quella ristretta del 1905. La nuova teoria passò test decisivi spiegando certe anomalie nel moto di Mercurio, ebbe un trionfo nel 1919 prevedendo esattamente la deflessione della luce nel campo gravitazionale del Sole [cfr. Dyson, Eddington e Davidson 1920]. Nel centenario della nascita di Einstein, l'evidenza a favore della teoria generale è ormai schiacciante e ad essa ha contribuito in modo notevole la tecnologia spaziale [cfr. Woolf 1979].

Colpisce, nella relatività generale, non tanto la predizione di sottili fenomeni, quanto il suo sorgere da esigenze estetiche lontane dall'evidenza empirica di tipo nozionistico e fenomenologico. Il punto di partenza rimane di una semplicità esemplare: l'uguaglianza tra massa inerziale e massa gravitazionale. Di qui si giunge al principio di equivalenza.

Raggiunta la formulazione finale della teoria nel 1916, scritte le equazioni del campo nella loro forma piú perspicua, toccò a Schwarzschild [1916] derivare la famosa soluzione a simmetria sferica, analoga al potenziale di Newton, che doveva avere un ruolo importantissimo nell'astrofisica relativistica [cfr. Misner, Thorne e Wheeler 1973]. Gli anni '20 videro la costruzione e la messa in opera di telescopi di grande apertura che permisero a Hubble [1929] di annunziare la legge di recessione delle galassie. Questi dati osservativi confluirono nell'impostazione relativistica della cosmologia, la recessione trovò una spiegazione naturale nel modello del *big bang* [cfr. Weinberg 1972; Misner, Thorne e Wheeler 1973].

La relatività, nata dal microcosmo delle particelle elementari, aveva finalmente aperto una finestra sull'universo e sulle nostre origini. Il lavoro non è ancora finito, i dati in nostro possesso sembrano confermare l'ipotesi del *big bang*; risulta ovvio che la cosmologia è ancora agl'inizi. Ai quesiti che essa ha posto daranno forse una risposta il nuovo telescopio spaziale, i radiotelescopi e tutta la tecnologia spaziale.

1. *Il principio di relatività.*

1.1. Lo spazio-tempo.

Si ammetterà in quanto segue che i fenomeni fisici si svolgano nell'ambito di un continuo, detto spazio-tempo, $\mathfrak{M}$. Gli elementi dello spazio-tempo si dicono eventi e sono caratterizzati da tre coordinate spaziali x, y, z e dal tempo t. Ogni fenomeno fisico si può rappresentare come una successione di eventi in $\mathfrak{M}$. Lo spazio-tempo contiene materia ed è compito della fisica studiarne le proprietà. Non si conosce alcuna giustificazione teorica aprioristica sulle cause di questa scelta di dimensioni per $\mathfrak{M}$; esistono invece continue prove empiriche di questo fatto, che inducono ad accettarlo come evidente.

Le coordinate x, y, z furono introdotte da Descartes e servono ad identificare i punti dello spazio ordinario. Si userà la notazione matematica $\mathbf{R}$ per la retta reale, $\mathbf{R}^2$ per il piano, $\mathbf{R}^3$ per lo spazio. Secondo la stessa nomenclatura, per lo spazio-tempo $\mathfrak{M}$ si dovrebbe usare $\mathbf{R}^4$. Tuttavia, con $\mathfrak{M}$ s'intenderà $\mathbf{R}^4$ dotato di una struttura metrica derivata dalla fisica e atta a rappresentare i fenomeni relativistici. Chiaramente una certa ipersuperficie $t=\text{cost}$ in $\mathfrak{M}$ può essere considerata come un particolare $\mathbf{R}^3$. Sotto questo profilo $\mathfrak{M}$ è una successione continua di $\mathbf{R}^3$ lungo la linea del tempo.

In $\mathbf{R}^3$ esistono infinite maniere diverse d'introdurre sistemi di coordinate in modo equivalente. Un sistema di coordinate presuppone infatti l'esistenza di un punto privilegiato, l'origine O, con coordinate $x=y=z=0$, e di tre assi ortogonali per O, rispettivamente $x=y=0$ (asse z), ecc. Nulla impedisce di porre O in un punto qualunque di $\mathbf{R}^3$ e di orientare gli assi intorno ad O in posizione qualunque. Tutti questi sistemi di riferimento sono ugualmente accettabili dal punto di vista fisico. È possibile valutare il grado di arbitrarietà nella scelta del sistema. Sia dato un sistema inerziale qualunque. Ogni altro sistema sarà caratterizzato dalla nuova posizione dell'origine O' (tre coordinate) e dall'inclinazione dei nuovi assi rispetto ai vecchi, per cui necessitano tre angoli, in totale dunque sei variabili.

Qual è la relazione tra diversi sistemi di coordinate? Essa discende dall'esistenza del concetto di distanza tra punti di $\mathbf{R}^3$. Siano x_1, y_1, z_1; x_2, y_2, z_2 due punti in $\mathbf{R}^3$. La loro distanza è definita dal teorema di Pitagora:

$$d^2=(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2. \tag{1.1}$$

In quanto saranno spesso considerate coppie di punti infinitamente vicini, si avranno dunque $x_2=x_1+dx$, $y_2=y_1+dy$, $z_2=z_1+dz$, ossia

$$d^2=dx^2+dy^2+dz^2. \tag{1.2}$$

Si fa riferimento a formule del tipo (1.1) e (1.2) come alla metrica di $\mathbf{R}^3$. Ora si può dimostrare che ogni relazione del tipo

$$x' = f(x,y,z) \qquad y' = g(x,y,z) \qquad z' = h(x,y,z) \tag{1.3}$$

tra vecchie coordinate x, y, z e nuove coordinate x', y', z', che implichi

$$dx^2 + dy^2 + dz^2 = (dx')^2 + (dy')^2 + (dz')^2, \tag{1.4}$$

è necessariamente lineare, e cioè del tipo

$$\begin{aligned} x' &= \mathcal{O}^{11}x + \mathcal{O}^{12}y + \mathcal{O}^{13}z + \xi^1 \\ y' &= \mathcal{O}^{21}x + \mathcal{O}^{22}y + \mathcal{O}^{23}z + \xi^2 \\ z' &= \mathcal{O}^{31}x + \mathcal{O}^{32}y + \mathcal{O}^{33}z + \xi^3 \end{aligned} \tag{1.5}$$

dove $\mathcal{O}^{ij}$, ξ^i sono costanti e formano rispettivamente una matrice $\mathcal{O}$ (3×3) e ξ (3×1). Anche $x = x^1$, $y = x^2$, $z = x^3$ formano una matrice

$$x = \left\| \begin{matrix} x^1 \\ x^2 \\ x^3 \end{matrix} \right\|,$$

per cui (1.5) si scrive anche

$$x' = \mathcal{O}x + \xi. \tag{1.6}$$

Inoltre $\mathcal{O}$ è una matrice ortogonale, ossia $\mathcal{O}^T\mathcal{O} = 1$ (1 matrice identità), dove $\mathcal{O}^T$ è la matrice trasposta di $\mathcal{O}$: $(\mathcal{O}^T)^{ij} = \mathcal{O}^{ji}$.

L'ortogonalità di $\mathcal{O}$ impone sei condizioni indipendenti sugli elementi della matrice. Bastano quindi tre parametri indipendenti per rappresentare le nove $\mathcal{O}^{ik}$, i cosiddetti angoli di Eulero usati nella teoria del rotatore (cfr. oltre, § 3.1). Si ha, come previsto,

$$\begin{aligned} d^2 &= dx^T\, dx \\ dx' &= \mathcal{O}\, dx \\ (d')^2 &= (dx')^T\, dx' = dx^T \mathcal{O}^T \mathcal{O}\, dx = dx^T\, dx = d^2. \end{aligned}$$

Finora è stata discussa la geometria di $\mathbf{R}^3$, e non quella di $\mathfrak{M}$. Si è trascurato di proposito il ruolo del tempo, perché necessita di un'analisi del concetto di sistema inerziale e costringe ad entrare nel vivo della trattazione relativistica. Il punto di partenza per lo studio della struttura di $\mathfrak{M}$ è la constatazione empirica dell'esistenza di infiniti sistemi di riferimento inerziali nello spazio-tempo.

Per chiarire il concetto di sistema inerziale occorre citare il principio d'inerzia di Galileo: «Ogni corpo persevera nel proprio stato di riposo o di movimento uniforme e rettilineo, a meno che non sia costretto a modificare tale stato da forze agenti su di esso».

Non è possibile accettare il principio d'inerzia senza un'analisi critica. Esso si riferisce infatti a una classe particolare di osservatori, detti inerziali. Se si fosse ad esempio racchiusi in una cabina soggetta a rotazione, apparirebbero forze centrifughe e di Coriolis che farebbero sembrare falso il principio. Analogamente, sulla Terra la rotazione diurna è evidenziabile da fenomeni, rilevabili con il metodo di Foucault, che dimostrano in ultima analisi una violazione del principio d'inerzia.

Quali sono dunque l'informazione e la portata effettiva di tale principio? Esso asserisce l'esistenza di una classe privilegiata di osservatori, quelli inerziali, e li distingue da quelli che non lo sono. La relatività ristretta conduce a relazioni precise tra osservazioni effettuate da diversi osservatori inerziali che mostrano strette analogie con la (1.5).

Prima di entrare nel formalismo matematico converrà soffermarsi su alcuni aspetti intuitivi del principio di relatività. Quando sia noto un sistema inerziale A, ci si aspetta che tutti gli altri si muovano di moto uniforme rispetto ad A, differendone anche per l'inclinazione degli assi. Se vale dunque il principio di Galileo in A, ci si aspetta che esso sia valido per un qualunque altro osservatore. Piú generalmente, il principio di relatività dice che «le leggi di natura devono scriversi in modo rigorosamente identico per tutti gli osservatori inerziali e non possono essere usate per distinguere o privilegiare uno di essi in particolare».

Sia nella meccanica di Newton sia in quella di Einstein vale il principio di relatività, che lega tra di loro le osservazioni di diversi sistemi inerziali. La forma matematica di questo legame è invece differente nei due casi. A basse velocità è praticamente impossibile distinguere la meccanica relativistica da quella newtoniana. A velocità prossime a quella della luce queste differenze sono rilevanti.

Sia la relatività di Galileo sia quella di Einstein sono intrinsecamente consistenti e risolvono in modo diverso il problema di collegare le osservazioni fatte da diversi sistemi inerziali. Non esiste alcun criterio intrinseco per decidere sulla validità o meno di ciascuna alternativa: pesanti paradossi scoperti a carico della relatività di Einstein derivano dall'introduzione nella teoria di elementi ad essa estranei. La decisione finale è giustificata in ultima analisi dai dati empirici.

L'apprendimento della teoria einsteiniana è legato al superamento di molti pregiudizi, di natura psicologica e soggettiva, che ne ritardarono per anni l'accettazione da parte degli elementi piú tradizionali della comunità scientifica. Particolarmente osteggiato fu l'abbandono del concetto di tempo universale, separato dalle coordinate, che, come diceva Newton, fluisce in modo uguale per tutti gli osservatori.

I vantaggi derivanti dal superamento di questi pregiudizi furono immensi. La relatività si è dimostrata capace di correlare in un formalismo semplice e scorrevole fenomeni apparentemente disparati ed ha introdotto una metodologia basata sull'unificazione dei concetti fisici il cui effetto è ancora evidente ed operante ad oltre settant'anni dalla scoperta.

1.2. Relatività einsteiniana e relatività galileiana.

La meccanica newtoniana rappresentò la fase culminante di un lungo processo storico inteso a correlare le osservazioni astronomiche. Il cosmo newtoniano consiste di un insieme di enti puntiformi chiamati corpi, dotati di un

attributo M (numero reale positivo) chiamato massa, e di una posizione $x \in \mathbf{R}^3$. L'insieme di tutte le posizioni successive $x(t)$ di un corpo nello spazio-tempo forma una linea, la traiettoria del corpo. Siano α, β indici etichettanti l'insieme dei corpi. Con il vettore

$$\mathbf{F}_{\alpha\beta} = -G\frac{\mathbf{x}_\alpha - \mathbf{x}_\beta}{d^3_{\alpha\beta}} M_\alpha M_\beta = -\mathbf{F}_{\beta\alpha} \tag{1.7}$$

s'intenderà la *forza* agente fra i corpi α, β, dove $\mathbf{x}_\alpha$ è la coordinata (in forma vettoriale) e M_α la massa del corpo α; $d_{\alpha\beta}$ è la distanza fra β e α, G la costante universale di Newton.

Il secondo principio della dinamica offre un'equazione differenziale del secondo ordine per il moto dei corpi

$$M_\alpha \frac{d^2\mathbf{x}_\alpha}{dt^2} = \sum_{\beta \neq \alpha} \mathbf{F}_{\alpha\beta} = \mathbf{F}_\alpha \tag{1.8}$$

che lega l'accelerazione $\boldsymbol{\delta}_\alpha \equiv (d^2\mathbf{x}_\alpha/dt^2)$ del corpo α alla forza totale $\mathbf{F}_\alpha$ di attrazione esercitata dagli altri corpi su α. Mediante la (1.7) e la (1.8) è possibile dedurre tutte le posizioni e velocità dei corpi per un tempo qualsiasi quando siano note all'istante iniziale, ad esempio $t = 0$, le quantità

$$\mathbf{x}_\alpha(0) \qquad \mathbf{v}_\alpha(0) = \left.\frac{d\mathbf{x}_\alpha(t)}{dt}\right|_{t=0}.$$

L'insieme delle (1.7) e (1.8) contiene in sintesi tutta la meccanica celeste che ha correlato in modo cosí straordinario il moto dei corpi entro il sistema solare.

Alla luce di quanto esposto in precedenza, si noti tuttavia:

I) Le (1.7) e (1.8) vanno intese come valevoli in un sistema inerziale A. Infatti se il corpo α viene allontanato da tutti gli altri, la $\mathbf{F}_\alpha$ si annulla e cosí pure l'accelerazione $\boldsymbol{\delta}_\alpha$, per cui il corpo si muove di moto uniforme come ci si aspetta per un sistema inerziale.

II) Sorge quindi spontanea la domanda: Quali sono allora gli altri infiniti sistemi inerziali? Da quanto detto essi sono tutti e solo i sistemi in moto uniforme rispetto a quello iniziale A.

Si indichino con $\mathbf{x}'_\alpha$ le nuove coordinate in un nuovo sistema inerziale B. Per semplicità si può ammettere che abbia assi paralleli a quelli di A. In generale si avrà

$$\mathbf{x}'_\alpha = \mathbf{x}_\alpha + \mathbf{u}t + \boldsymbol{\xi} \qquad t' = t + t_0 \tag{1.9}$$

dove $\mathbf{u}$, $\boldsymbol{\xi}$, t_0 sono costanti. Ma chiaramente $\mathbf{x}'_\alpha - \mathbf{x}'_\beta = \mathbf{x}_\alpha - \mathbf{x}_\beta$, $d'_{\alpha\beta} = d_{\alpha\beta}$, $dt = dt'$, dato che $\mathbf{u}$ non dipende da α; per cui $\mathbf{F}'_{\alpha\beta} = \mathbf{F}_{\alpha\beta}$, $\mathbf{F}'_\alpha = \mathbf{F}_\alpha$. Inoltre

$$\begin{aligned} \frac{d\mathbf{x}'_\alpha}{dt'} &= \frac{d\mathbf{x}'_\alpha}{dt} = \mathbf{v}_\alpha + \mathbf{u} = \mathbf{v}'_\alpha \\ \frac{d^2\mathbf{x}'_\alpha}{d(t')^2} &= \frac{d^2\mathbf{x}_\alpha}{dt^2} = \boldsymbol{\delta}_\alpha = \boldsymbol{\delta}'_\alpha. \end{aligned} \tag{1.10}$$

Quindi se le (1.7) e (1.8) sono verificate nel sistema A, lo sono anche nel sistema B. Se dunque il cosmo fosse popolato esclusivamente da corpi gravitanti newtoniani, non sarebbe possibile, dalla validità o meno delle (1.7) e (1.8), dedurre alcunché sullo stato di moto assoluto di un sistema, dato che esse si scrivono esattamente nello stesso modo. In questo senso la meccanica di Newton è una teoria relativistica dove le (1.9) legano le descrizioni dello stesso evento in $\mathfrak{M}$ ottenute rispettivamente in B e A.

Le (1.9) sono trasformazioni lineari nelle coordinate $\mathbf{x}$, t di un evento in $\mathfrak{M}$ e sono molto simili alle (1.5) adottate per $\mathbf{R}^3$. In generale, se si ammettono assi inclinati per A e B, si avrà:

$$\begin{aligned} x'^i &= \sum_{k=1}^{3} \mathcal{O}^{ik} x^k + u^i t + \xi^i \qquad (i = 1, 2, 3) \\ t' &= t + t_0. \end{aligned} \tag{1.11}$$

Si farà riferimento alle (1.11) come alla piú generica «trasformazione di Galileo» tra le coordinate x'^i, t' in B e x^i, t in A. Essa contiene in totale dieci parametri (tre in ξ^i, tre in u^i, tre in $\mathcal{O}^{ik}$, e uno in t_0). Come si vedrà, questo conteggio rimane invariato nelle trasformazioni di Lorentz che valgono per la teoria einsteiniana. Non esiste alcun dubbio che le (1.7) e le (1.8), insieme alle (1.11), rappresentino con grande semplicità e accuratezza una vasta messe di osservazioni.

Le (1.11) sono legate alla (1.10) che (trascurando come nella (1.11) l'indice del corpo) può essere riscritta come

$$\mathbf{v}' = \mathbf{v} + \mathbf{u}, \tag{1.12}$$

ossia come il teorema di addizione delle velocità. Questo teorema è dunque legato alle (1.11). Per quanto sia di grande semplicità e potenza intuitiva, il teorema (1.12) è risultato empiricamente il punto piú debole della relatività galileiana; da esso ha preso le mosse la critica einsteiniana.

La crisi giunse per il tramite della costruzione finale, operata da Maxwell, della teoria elettromagnetica della luce. La teoria di Maxwell introduce nel cosmo, oltre ai corpi gravitanti di Newton, due nuove entità, il campo elettromagnetico (formato dalla coppia $\mathbf{E}$, $\mathbf{B}$) e un nuovo attributo, la carica elettrica. Secondo Maxwell si hanno le equazioni omogenee

$$\operatorname{rot}\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} \qquad \operatorname{div}\mathbf{B} = 0 \tag{1.13}$$

e quelle non omogenee

$$\operatorname{rot}\mathbf{B} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{c}\mathbf{j} \qquad \operatorname{div}\mathbf{E} = 4\pi\rho \tag{1.14}$$

dove ρ, $\mathbf{j}$ rappresentano rispettivamente la densità di carica elettrica e la corrente elettrica per unità di superficie. Le (1.13) e (1.14) possiedono nel vuoto (cioè dove $\rho = \mathbf{j} = 0$) una classe di soluzioni a carattere oscillante (in $\mathbf{E}$, $\mathbf{B}$), che si pro-

paga con velocità $c=(299\,792{,}5\pm 0{,}3)$ km/sec, e che va identificata (nell'appropriata banda di frequenze) con le onde luminose previste dalla teoria ondulatoria della luce di Huygens.

La capacità delle equazioni (1.13) e (1.14) di predire una larga messe di dati sperimentali superò presto quella seppur rilevante della teoria di Newton. Esiste tuttavia una macroscopica differenza fra le (1.13) e (1.14), e le (1.7) e (1.8), in quanto, se si accetta la (1.11), e quindi anche la (1.12), le equazioni di Maxwell non possono che valere in un solo sistema di riferimento, in altre parole, esse *non* rispettano la relatività galileiana. Quest'asserzione segue dalla facile constatazione che le onde luminose si propagano appunto con la velocità c della luce. Se quindi si ammette la validità delle (1.13) e (1.14) in un sistema di riferimento inerziale A, esse non possono essere valide in B, perché in B la velocità dell'onda luminosa non sarebbe piú $c\mathbf{i}$ ($\mathbf{i}$ versore dell'asse x), dovendosi aggiungere vettorialmente ad essa la velocità $\mathbf{v}$ di B rispetto ad A, secondo la (1.12). Chiaramente, per $\mathbf{v}$ generico, $|c\mathbf{i}+\mathbf{v}|$ è diverso da c.

All'inizio del secolo si pensava che la luce si propagasse in un mezzo elastico (l'etere) analogamente alla propagazione del suono nell'aria. Il sistema inerziale A a riposo rispetto all'etere era quello in cui erano valide le equazioni di Maxwell, mentre gli altri avrebbero dovuto mostrare deviazioni dovute alla diversa velocità di propagazione della luce nelle diverse direzioni. Secondo questo punto di vista si mossero Michelson e Morley per misurare direttamente il «vento d'etere» e l'anisotropia della propagazione della luce: l'esperimento dette un clamoroso risultato negativo.

Secondo Einstein, il fallimento dell'esperimento di Michelson e di altri analoghi che si proponevano di misurare lo stato assoluto di moto di un sistema inerziale era prevedibile se si adottava l'idea secondo cui l'esistenza di un principio di relatività era piú importante e desiderabile del mantenimento delle (1.11) e (1.12).

Secondo Einstein, dunque, neppure le equazioni di Maxwell dovrebbero permettere d'identificare un sistema inerziale privilegiato e le (1.11) e (1.12) devono venire modificate per tener conto di questa condizione, secondo quanto già indicato dal fisico olandese Lorentz. La critica alle (1.11) e (1.12) coinvolge indirettamente una critica alle (1.7) e (1.8), che in un certo senso le implicano, in quanto le possiedono come gruppo d'invarianza.

Tenendo conto dell'enorme prestigio della meccanica newtoniana, già vastamente popolarizzata e ispiratrice di splendidi lavori di meccanica celeste, oltreché della sintesi operata da Lagrange, Hamilton e Jacobi, si può facilmente comprendere come l'idea di modificarla, per mantenere in vita un'astrusa equazione di campo comprensibile solo da pochi iniziati, abbia incontrato non poche difficoltà. Solo un profondo senso della realtà fisica poteva condurre a una scelta che appariva, ai piú, folle e inutilmente audace. E questo convincimento non poteva che nascere dalla constatazione che la precisione e la vastità dei risultati predetti dalle equazioni di Maxwell già superava quella, sia pur splendida, della sintesi di Newton, e che le modificazioni indotte in questa dalla relatività einsteiniana ne avrebbero salvate le parti essenziali, aprendo in cambio orizzonti nuovi e sconfinati; per queste ragioni fu operata la scelta di abbandonare le (1.11) e (1.12) e di postulare invece:

I) esistenza di infiniti sistemi inerziali;

II) allargamento del postulato di relatività, sino ad includere, oltre ai corpi materiali, anche la luce e cioè le equazioni di Maxwell. In particolare, si assume che queste si scrivano nella stessa forma (1.13) e (1.14) in tutti i sistemi inerziali e che quindi la luce si propaghi con velocità c (costante universale pari a circa 300 000 km/sec) in tutte le direzioni e in tutti i sistemi inerziali.

Il postulato II) contraddice apertamente la (1.12), secondo la quale, inseguendo la luce con velocità $\mathbf{u}$, la si vedrebbe allontanarsi con velocità $\mathbf{c}-\mathbf{u}$, mentre fuggendola con velocità $\mathbf{v}$ la si vedrebbe muoversi con velocità $\mathbf{c}+\mathbf{v}$, con risultati intermedi per le altre direzioni.

Si abbandoneranno quindi le (1.11) e (1.12) e, nella forma presente, le (1.7) e (1.8), considerandole approssimazioni valide a basse velocità (rispetto a c) della nuova teoria, ipotesi che risulterà confermata a posteriori. Dal postulato II) verranno derivate le nuove leggi di trasformazione (dette di Lorentz) che sostituiscono le (1.11) e (1.12) e che stanno alla base della fisica moderna.

2. *Cinematica relativistica.*

2.1. Le trasformazioni di Lorentz: derivazione.

Nello sviluppare le conseguenze del postulato II), ci si limiterà a casi in cui la configurazione relativa degli osservatori A, B risulta particolarmente semplice. Si assumerà infatti che A, B abbiano assi paralleli e origine comune al tempo $t=0$. Inoltre la velocità v relativa sarà diretta lungo l'asse x.

Ci si propone di modificare le (1.11) e (1.12) in modo tale che un fronte d'onda che si propaga con velocità c nel sistema A mantenga tale velocità nel sistema B. Per quanto detto, la (1.11) dovrebbe ridursi alle

$$x'=x+vt \qquad y'=y \qquad z'=z \qquad t'=t \tag{2.1}$$

che risultano incompatibili con la costanza della velocità della luce. Al momento ci si occuperà di modificazioni della (1.11) di tipo lineare omogeneo,

$$x'=\Lambda_1{}^1x+\Lambda_1{}^4t \qquad y'=\Lambda_2{}^2y \qquad z'=\Lambda_3{}^3z \qquad t'=\Lambda_4{}^1x+\Lambda_4{}^4t, \tag{2.2}$$

rimandando al § 3 lo studio delle trasformazioni generiche. Un punto qualsiasi del fronte d'onda centrato in O obbedisce all'equazione

$$x^2+y^2+z^2=c^2t^2 \tag{2.3}$$

che deve valere anche per x', y', z', t'. Ne segue:

$$\begin{aligned}(x')^2+(y')^2+(z')^2-c^2(t')^2=&[(\Lambda_1{}^1)^2-c^2(\Lambda_4{}^1)^2]x^2+(\Lambda_2{}^2)^2y^2+\\&+(\Lambda_3{}^3)^2z^2+2(\Lambda_1{}^1\Lambda_1{}^4-c^2\Lambda_4{}^1\Lambda_4{}^4)xt+[(\Lambda_1{}^4)^2-c^2(\Lambda_4{}^4)^2]t^2=0.\end{aligned} \tag{2.4}$$

Questa relazione deve ridursi a un multiplo della (2.3), $x^2+y^2+z^2=c^2t^2$, per cui si ha:

$$(2.5)\qquad \begin{aligned} &(\Lambda_1{}^1)^2-c^2(\Lambda_4{}^1)^2=\lambda \\ &\Lambda_1{}^1\Lambda_1{}^4-c^2\Lambda_4{}^1\Lambda_4{}^4=0 \\ &(\Lambda_1{}^4)^2-c^2(\Lambda_4{}^4)^2=-\lambda c^2 \end{aligned} \qquad \begin{aligned} &(\Lambda_2{}^2)^2=\lambda>0 \\ &(\Lambda_3{}^3)^2=\lambda>0. \end{aligned}$$

Data una soluzione delle (2.5), se ne ottiene un'altra cambiando dei segni secondo la

$$(2.6)\qquad \Lambda_a{}^i\to(-1)^{\varepsilon_a}\Lambda_a{}^i,$$

dove le ε_a sono arbitrarie. Si può dunque supporre $\Lambda_a{}^a>0$, potendosi ricavare gli altri casi mediante semplici cambiamenti di segno. Ponendo

$$(2.7)\qquad \Lambda_1{}^4=v\Lambda_4{}^4,$$

si ottiene dalle (2.5) $v\Lambda_1{}^1=c^2\Lambda_4{}^1$ e

$$(2.8)\qquad \Lambda_1{}^1=\Lambda_4{}^4=\frac{\sqrt{\lambda}}{\sqrt{1-\beta^2}}\qquad \Lambda_4{}^1=\frac{v\sqrt{\lambda}}{c^2\sqrt{1-\beta^2}}$$

ove, seguendo la consuetudine, si ponga $\beta=v/c$. L'origine O di A ($x=y=z=0$) appare muoversi in B secondo la

$$(2.9)\qquad x'=\Lambda_1{}^4(\Lambda_4{}^4)^{-1}t'=vt'$$

per cui v va interpretata come la velocità relativa degli osservatori A e B. Affinché la (2.8) abbia senso, occorre che $|\beta|<1$. Dalle (2.7) e (2.8) segue:

$$(2.10)\qquad \begin{aligned} t' &= \frac{\sqrt{\lambda}}{\sqrt{1-\beta^2}}\left(\frac{v}{c^2}x+t\right) & x' &= \frac{\sqrt{\lambda}}{\sqrt{1-\beta^2}}(x+vt) \\ y' &= \sqrt{\lambda}y & z' &= \sqrt{\lambda}z. \end{aligned}$$

La trasformazione deve essere completamente determinata da v, quindi λ deve essere una funzione di v: $\lambda=\lambda(v)$. La scelta dell'asse x è puramente convenzionale e λ deve dipendere in generale dal modulo della velocità relativa di A e di B, per cui $\lambda(v)=\lambda(-v)$. Invertendo la (2.10) si trova

$$(2.11)\qquad \begin{aligned} t &= \frac{(\sqrt{\lambda})^{-1}}{\sqrt{1-\beta^2}}\left(-\frac{v}{c^2}x'+t'\right) & x &= \frac{(\sqrt{\lambda})^{-1}}{\sqrt{1-\beta^2}}(x'-vt') \\ y &= (\sqrt{\lambda})^{-1}y' & z &= (\sqrt{\lambda})^{-1}z' \end{aligned}$$

che differisce dalla (2.10) per l'inversione dei ruoli di x', x, ecc., per lo scambio di v in $-v$ e di $\lambda(v)$ in $(\lambda(v))^{-1}$. Dunque $\lambda(-v)=(\lambda(v))^{-1}$, ossia $\lambda=1$. Si ottiene la ben nota trasformazione di Lorentz

$$(2.12)\qquad t'=\frac{t+\frac{v}{c^2}x}{\sqrt{1-\beta^2}}\qquad x'=\frac{x+vt}{\sqrt{1-\beta^2}}\qquad y'=y\qquad z'=z.$$

Operando su questa con le (2.6) e con le rotazioni, si ottiene la soluzione piú generale, che verrà discussa in seguito.

2.2. Proprietà delle trasformazioni di Lorentz.

In quanto segue verranno esaminate in dettaglio le proprietà delle (2.12).

2.2.1. *Metrica.* Per quanto detto si ha anzitutto

$$x^2+y^2+z^2-c^2t^2=(x')^2+(y')^2+(z')^2-c^2(t')^2.$$

Analogamente, per i differenziali si ha:

$$(2.13)\qquad du^2\equiv c^2dt^2-\sum_{i=1}^{3}(dx^i)^2=c^2(dt')^2-\sum_{i=1}^{3}(dx'^i)^2.$$

La (2.13) esprime ancora una volta la costanza della velocità della luce.

2.2.2. *Trasformazioni di Lorentz.* Introducendo la notazione $x^0=ct$, $x'^0=ct'$, che sarà usata estesamente in seguito, si ha:

$$(2.14)\qquad x'^0=\frac{x^0+\beta x^1}{\sqrt{1-\beta^2}}\qquad x'^1=\frac{x^1+\beta x^0}{\sqrt{1-\beta^2}}\qquad x'^2=x^2\qquad x'^3=x^3$$

$$(2.15)\qquad -(x'^0)^2+(x'^1)^2=-(x^0)^2+(x^1)^2,$$

che non contiene piú esplicitamente la velocità della luce c. La velocità appare in unità adimensionate β.

2.2.3. *Rotazioni nel piano* x^0, x^1. Sia $\beta=\operatorname{tgh}\vartheta$, dove ϑ prende il nome di rapidità e $\operatorname{tgh}\vartheta=(e^{\vartheta}-e^{-\vartheta})/(e^{\vartheta}+e^{-\vartheta})$ è la cosiddetta tangente iperbolica di ϑ. Da questa definizione seguono le relazioni

$$(2.16)\qquad \cosh\vartheta=\frac{1}{\sqrt{1-\beta^2}}\qquad \sinh\vartheta=\frac{\beta}{\sqrt{1-\beta^2}},$$

per cui le (2.14) ora si possono scrivere come

$$(2.17)\qquad \begin{aligned} &x'^0=\sinh\vartheta\, x^1+\cosh\vartheta\, x^0 \qquad x'^1=\cosh\vartheta\, x^1+\sinh\vartheta\, x^0 \\ &x'^2=x^2 \qquad x'^3=x^3. \end{aligned}$$

La rotazione ordinaria nel piano euclideo x^2, x^3 dell'angolo φ si scrive invece

$$(2.18)\qquad \begin{aligned} &x'^0=x^0 \qquad x'^1=x^1 \\ &x'^2=\cos\varphi\, x^2+\sin\varphi\, x^3 \qquad x'^3=-\sin\varphi\, x^2+\cos\varphi\, x^3 \end{aligned}$$

e implica

$$(2.19)\qquad (x^2)^2+(x^3)^2=(x'^2)^2+(x'^3)^2;$$

come appare da un semplice confronto, la (2.18) e la (2.19) differiscono dalle (2.17) e (2.15) per lo scambio

$$\begin{aligned} &\varphi \to i\vartheta \\ &x^2 \to x^1 \qquad x^3 \to -ix^0. \end{aligned} \tag{2.20}$$

Questa coincidenza può essere usata come guida formale per anticipare il risultato di calcoli relativistici partendo da altri ben noti per le rotazioni ordinarie. Grosso modo, il tempo può essere considerato come una variabile spaziale cui si attribuiscano valori immaginari. L'uso incontrollato di quest'analogia non è scevro da pericoli ed è fonte di speculazioni metafisiche che esulano dalle scienze naturali. Tuttavia, usato con le dovute cautele, si è rivelato un procedimento di alto valore in molte questioni di teoria dei campi e di meccanica statistica.

2.2.4. *Proprietà gruppali.* Il prodotto di due rotazioni ordinarie di angoli φ_1, φ_2 è un'altra rotazione di angolo $\varphi_1+\varphi_2$. Siano invece date le trasformazioni di Lorentz:

$$\begin{aligned} &t' = \frac{t+\dfrac{v_1 x}{c^2}}{\sqrt{1-\beta_1^2}} \qquad x' = \frac{x+v_1 t}{\sqrt{1-\beta_1^2}} \qquad y'=y \qquad z'=z \\ &t'' = \frac{t'+\dfrac{v_2 x'}{c^2}}{\sqrt{1-\beta_2^2}} \qquad x'' = \frac{x'+v_2 t}{\sqrt{1-\beta_2^2}} \qquad y''=y' \qquad z''=z'. \end{aligned} \tag{2.21}$$

Eliminando x', t', si ottiene quella prodotto

$$t'' = \frac{t+\dfrac{v_3 x}{c^2}}{\sqrt{1-\beta_3^2}} \qquad x'' = \frac{x+v_3 t}{\sqrt{1-\beta_3^2}} \qquad y''=y \qquad z''=z \tag{2.22}$$

dove la velocità v_3 è data dalla formula di composizione

$$v_3 = \frac{v_1+v_2}{1+\dfrac{v_1 v_2}{c^2}} \tag{2.23}$$

che sostituisce il vecchio teorema di addizione delle velocità:

$$v_3 = v_1 + v_2. \tag{2.24}$$

Ponendo

$$\beta_1 = \operatorname{tgh}\vartheta_1 \qquad \beta_2 = \operatorname{tgh}\vartheta_2, \tag{2.25}$$

si ottiene esattamente $\beta_3 = \operatorname{tgh}(\vartheta_1+\vartheta_2)$. Nelle rapidità ϑ_1, ϑ_2 la composizione è dunque additiva come negli angoli φ_1, φ_2. La (2.23) dà gli stessi risultati della (2.24) per $v_1, v_2 \ll c$. Se $v_1 = c$, si ha sempre $v_3 = c$, qualunque sia il valore di v_2; si riottiene la costanza della velocità della luce. Per valori intermedi si ottengono risultati poco familiari. Se $v_1 = v_2 = c/2$ si ottiene $v_3 = (4/5)c$, e se $v_1 = v_2 = 0{,}9c$ si ha $v_3 = 1{,}8c/1{,}81 = c-(c/181) < c$, per cui, anche componendo velocità assai prossime a c, non è possibile superare il muro della luce.

2.2.5. *Simultaneità.* Siano $x_1 = t_1 = t_2 = 0$, $x_2 = L$ due eventi simultanei in A (si supporrà in quanto segue $y = y' = z = z' = 0$). Si ha

$$\begin{aligned} &t_1' = \frac{t_1+\dfrac{v x_1}{c^2}}{\sqrt{1-\beta^2}} = 0 \qquad x_1' = \frac{x_1+v t_1}{\sqrt{1-\beta^2}} \\ &t_2' = \frac{vL}{c^2}\frac{1}{\sqrt{1-\beta^2}} \neq t_1' \qquad x_2' = \frac{L}{\sqrt{1-\beta^2}} \end{aligned}$$

per cui i due eventi non sono simultanei in B.

L'idea di una simultaneità relativa fu considerata palesemente assurda da molti oppositori della relatività. Come dice Einstein [1922], l'opposizione è venuta da filosofi che hanno spostato certi concetti fondamentali dal dominio dell'empirismo, e quindi dal nostro controllo, alle altezze intangibili dell'«a priori». L'effetto è reale ma ridicolmente piccolo a basse velocità. Per un aeroplano si ha tipicamente $v \sim 300$ m/sec, ossia $\beta \sim 10^{-6}$; per $L = 3$ km si trova $t_2 - t_1 \sim 10^{-11}$ sec, difficilmente rilevabile anche se di altissimo valore concettuale.

2.2.6. *Contrazione delle lunghezze.* Sia in B una sbarra ferma di lunghezza L i cui estremi siano quindi eventi: $x_1' = 0$, t_1' qualsiasi, $x_2' = L$; t_2' qualsiasi. La sbarra viene osservata da A ed appare in movimento con velocità v. Per definizione, la sua lunghezza L' è data da $L' = x_2(t) - x_1(t)$, essendo $(x_1(t)+vt)/\sqrt{1-\beta^2} = 0$, $(x_2(t)+vt)/\sqrt{1-\beta^2} = L$, dove appare la stessa variabile t in ambedue i termini. Risolvendo questo sistema, si ottiene

$$L' = L\sqrt{1-\beta^2}, \tag{2.26}$$

che non dipende da t e che dà una lunghezza contratta del fattore $\sqrt{1-\beta^2}$ rispetto a quella della sbarra immobile.

3. *Lo spazio-tempo di Minkowski.*

3.1. Il gruppo di Lorentz.

Si noti che, mentre la (2.14) privilegia la coppia x^0, x^1, la (2.13) è completamente simmetrica nel ruolo di x^1, x^2, x^3. Quindi anche

$$x'^0 = \frac{x^0+\beta x^2}{\sqrt{1-\beta^2}} \qquad x'^1 = x^1 \qquad x'^2 = \frac{x^2+\beta x^0}{\sqrt{1-\beta^2}} \qquad x'^3 = x^3 \tag{3.1}$$

$$x'^0 = \frac{x^0+\beta x^3}{\sqrt{1-\beta^2}} \qquad x'^1 = x^1 \qquad x'^2 = x^2 \qquad x'^3 = \frac{x^3+\beta x^0}{\sqrt{1-\beta^2}} \tag{3.2}$$

sono trasformazioni di Lorentz. Effettuando successivamente le trasformazioni (2.14), (3.1), (3.2) in vario ordine, si arriverebbe a costruire una trasformazione di Lorentz generale in cui A e B si muovono con velocità relativa qualunque. È poco saggio tuttavia cercare di costruire questa trasformazione senza sapere di piú sulle proprietà generali delle trasformazioni di Lorentz. Per giungere a una discussione piú esauriente, occorre ampliare il formalismo, anticipando alcune tecniche di calcolo tensoriale.

Occorre quindi introdurre la matrice g (4×4) di elementi

$$(3.3)\qquad \begin{aligned} g_{\mu\nu} &= 0 && \text{se } \mu\neq\nu \\ g_{\mu\nu} &= -1 && \text{se } \mu=\nu=0 \\ g_{\mu\nu} &= +1 && \text{se } \mu=\nu=i \end{aligned}$$

e la matrice g^{-1}: $g^{\mu\nu}=g_{\mu\nu}$. L'uso di una matrice $g^{\mu\nu}$ con elementi uguali a $g_{\mu\nu}$ sembra uno spreco inutile di definizioni. Si vedrà tuttavia che 1) in relatività generale si usa $g^{\mu\nu}\neq g_{\mu\nu}$; 2) il ruolo di $g^{\mu\nu}$ e $g_{\mu\nu}$ è diverso nello sviluppo del calcolo tensoriale e ambedue le matrici sono necessarie per rispettare le regole sintattiche del calcolo.

Sia nelle $g_{\mu\nu}$, $g^{\mu\nu}$, sia nelle x^α, appaiono indici greci che prendono sempre i valori 0, 1, 2, 3. In quanto segue questa convenzione sarà mantenuta. Volendo usare un indice di tipo spaziale, che corre cioè nell'insieme 1, 2, 3, si adotterà una lettera latina i, j, k, ecc. Usando queste convenzioni, si vede che du^2 definito in (2.13) può essere scritto

$$(3.4)\qquad du^2 = -\sum_{\alpha,\beta=0}^{3} g_{\alpha\beta}\, dx^\alpha\, dx^\beta = -\sum_{\alpha,\beta=0}^{3} g_{\alpha\beta}\, dx'^\alpha\, dx'^\beta,$$

dove la somma è eseguita sugli indici α, β. In ciascun sommando, questi indici appaiono due volte; una volta come suffisso, e cioè in basso (ad esempio α in $g_{\alpha\beta}$), e un'altra volta in alto (ad esempio α in dx^α). Si ha nella (3.4) il primo esempio di formula tensoriale. Anche se verrà sviluppata piú avanti la tecnica del calcolo tensoriale, si possono già da questo paragrafo introdurre alcune convenzioni che semplificano notevolmente i calcoli e le notazioni, e che saranno usate estesamente trattando con i tensori. Si ha dunque:

1) In quanto segue verranno manipolati oggetti formati da un simbolo (g, dx, ecc.) accompagnato da uno o piú indici greci, ad esempio $g_{\alpha\beta}$, dx^α.
2) Saranno considerati sommandi formati da un prodotto dei simboli descritti in 1) con la restrizione che ciascun indice può apparire al massimo due volte; se un indice appare due volte, deve essere una volta in alto e una in basso.
3) Se un indice appare due volte, si intende sommato su tutto l'intervallo 0, 1, 2, 3, e si dice indice *muto*. Se un indice appare una volta sola, non è quindi sommato ed è detto *libero*.

Le regole 1), 2), 3) creano quindi un abbozzo di sintassi. Ad esempio, l'espressione $g_{\alpha\alpha}$ è «sintatticamente» inaccettabile, e cosí pure $\Gamma_\alpha{}^{\alpha\alpha}$ oppure $g_{\alpha\beta}A_\beta$.

Invece l'espressione $-g_{\alpha\beta}\, dx^\alpha\, dx^\beta$ deve intendersi uguale a du^2 secondo la (3.4). Dato che la sommatoria $\sum_{\alpha,\beta=0}^{3}$ in (3.4) è in accordo con la convenzione 3), la si può sopprimere dalla formula restando automaticamente intesa.

Considerando le matrici rettangolari

$$(3.5)\qquad x=\begin{Vmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{Vmatrix} \qquad x^T=\|x^0, x^1, x^2, x^3\| \qquad dx=\begin{Vmatrix} dx^0 \\ dx^1 \\ dx^2 \\ dx^3 \end{Vmatrix} \qquad \text{ecc.}$$

e usando il calcolo matriciale (si veda l'articolo «Invariante» in questa stessa *Enciclopedia*, VII, p. 901), si può scrivere

$$(3.6)\qquad -du^2 = g_{\alpha\beta}\, dx^\alpha\, dx^\beta = dx^T g\, dx.$$

Inoltre, sia la (2.14), che le (3.1) e (3.2) possono essere poste nella forma

$$(3.7)\qquad x'^\alpha = \Lambda^\alpha{}_\beta x^\beta \qquad \text{ovvero} \quad x'=\Lambda x,$$

dove Λ è una matrice quadrata di elementi $\Lambda^\alpha{}_\beta$ data da

$$(3.8)\qquad \Lambda^\alpha{}_\beta = \begin{Vmatrix} \frac{1}{\sqrt{1-\beta^2}} & 0 & \frac{\beta}{\sqrt{1-\beta^2}} & 0 \\ 0 & 1 & 0 & 0 \\ \frac{\beta}{\sqrt{1-\beta^2}} & 0 & \frac{1}{\sqrt{1-\beta^2}} & 0 \\ 0 & 0 & 0 & 1 \end{Vmatrix} \qquad \Lambda^\alpha{}_\beta = \begin{Vmatrix} \frac{1}{\sqrt{1-\beta^2}} & 0 & 0 & \frac{\beta}{\sqrt{1-\beta^2}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{\beta}{\sqrt{1-\beta^2}} & 0 & 0 & \frac{1}{\sqrt{1-\beta^2}} \end{Vmatrix}$$

nei casi (3.1) e (3.2) rispettivamente, e da

$$(3.9)\qquad \Lambda^{\alpha}{}_{\beta}=\begin{Vmatrix} \frac{1}{\sqrt{1-\beta^2}} & \frac{\beta}{\sqrt{1-\beta^2}} & 0 & 0 \\ \frac{\beta}{\sqrt{1-\beta^2}} & \frac{1}{\sqrt{1-\beta^2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{Vmatrix}$$

nel caso (2.14). Nelle (3.8), (3.9), l'indice α varia da 0 a 3 e corrisponde alle righe delle matrici ordinate dall'alto in basso; analogamente, β distingue le colonne, ordinate da sinistra a destra. È facile controllare che la proprietà di $\Lambda^{\alpha}{}_{\beta}$ di rappresentare una trasformazione di Lorentz assume una forma molto compatta usando le convenzioni adottate, oppure la notazione matriciale. In tutti i casi (3.8) e (3.9) si ha infatti

$$(3.10)\qquad g_{\alpha\beta}\Lambda^{\alpha}{}_{\nu}\Lambda^{\beta}{}_{\mu}=g_{\nu\mu}\qquad \Lambda^T g\Lambda=g;$$

queste relazioni risultano essere identicamente soddisfatte dalle (3.8) e (3.9) e implicano che, se

$$(3.11)\qquad dx'=\Lambda\, dx,$$

allora

$$(3.12)\qquad -(du')^2=dx'^T g\, dx'=dx^T g\, dx=-du^2;$$

infatti

$$dx'^T g\, dx'=dx^T \Lambda^T g\Lambda\, dx=dx^T g\, dx.$$

Si possono quindi generalizzare le formule (3.8) e (3.9) introducendo la definizione:

DEFINIZIONE 1. *Si dirà che $\Lambda^{\alpha}{}_{\beta}$ è una matrice di Lorentz se soddisfa alla condizione* (3.10). *L'insieme di tutte le matrici di Lorentz sarà denotato con* $O(3,1)$ *e chiamato gruppo di Lorentz* (per la definizione astratta di gruppo e per una introduzione ai gruppi classici, si vedano in questa stessa *Enciclopedia* gli articoli «Applicazioni», I, p. 718, e «Invariante», VII, pp. 925 sgg.). *Inoltre, trasformazione di Lorentz (omogenea) è una trasformazione del tipo $x'=\Lambda x$, dove $\Lambda\in O(3,1)$.*

DEFINIZIONE 2. *La trasformazione P di Poincaré (talvolta detta anche trasformazione di Lorentz non omogenea) è del tipo $x'=\Lambda x+\Xi$, dove $\Lambda\in O(3,1)$ e Ξ è una matrice reale 4×1 di elementi ξ^0, ξ^1, ξ^2, ξ^3. Gli elementi P possono quindi scriversi come coppie* (Λ, Ξ), *con identità* (1, 0) *e legge di composizione* $(\Lambda_1, \Xi_1)\cdot(\Lambda_2, \Xi_2)=(\Lambda_1\Lambda_2, \Lambda_1\Xi_2+\Xi_1)$.

P ha dieci parametri e l'insieme di queste trasformazioni forma gruppo. Nella meccanica quantistica lo studio delle rappresentazioni unitarie irriducibili di questo gruppo conduce alla importante classificazione delle particelle elementari secondo la loro massa e il loro spin.

Chiaramente, una trasformazione di Lorentz è un caso particolare di quella di Poincaré. In ogni modo anche per *P* si ha $dx'=\Lambda\, dx$. Per quanto segue, la traslazione Ξ non necessita di una particolare elaborazione, e sarà arbitraria.

Per mezzo del calcolo matriciale, si possono dedurre le proprietà generali delle $\Lambda\in O(3,1)$. Dato che tutte le matrici in (3.10) sono quadrate, si ha: $\det\Lambda=\pm 1$. Gli esempi (3.8) e (3.9) hanno $\det\Lambda=+1$. Il sottoinsieme di $O(3,1)$ formato dalle Λ con $\det\Lambda=+1$ si denota con $SO(3,1)$. Se $\det\Lambda=+1$, si dirà che Λ è propria, altrimenti che è impropria.

È facile verificare che (3.10) implica anche $(\Lambda^0{}_0)^2\geq 1$, ossia $\Lambda^0{}_0\leq -1$, oppure $\Lambda^0{}_0\geq +1$. Le $\Lambda\in O(3,1)$ per le quali $\Lambda^0{}_0\geq 1$ ($\Lambda^0{}_0\leq -1$) si dicono ortocrone (anticrone) e formano il sottoinsieme $O^+(3,1)$ ($O^-(3,1)$). Chiaramente: $O^+(3,1)\cup O^-(3,1)=O(3,1)$. Si scrive inoltre: $SO^{\pm}(3,1)=O^{\pm}(3,1)\cap SO(3,1)$. Vale la classificazione:

		Segno di $\det\Lambda$: +	Segno di $\det\Lambda$: −
Segno di $\Lambda^0{}_0$	+	$SO^+(3,1)$	Improprie ortocrone
	−	$SO^-(3,1)$	Improprie anticrone

È facile dimostrare che $O(3,1)$, $SO(3,1)$, $O^+(3,1)$, $SO^+(3,1)$ sono gruppi, mentre $SO^-(3,1)$ e le Λ improprie non sono gruppi. Date le matrici di $SO^+(3,1)$, è facile costruire le rimanenti. Una Λ impropria ortocrona è data da

$$(3.13)\qquad \Lambda_1=\begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{Vmatrix}.$$

Una $\Lambda\in SO^-(3,1)$ è

$$(3.14)\qquad \Lambda_2=\begin{Vmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{Vmatrix}$$

e una Λ impropria anticrona

$$(3.15)\qquad \Lambda_1\Lambda_2=\begin{Vmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{Vmatrix}.$$

Chiaramente l'insieme $\Lambda_1\times SO^+(3,1)$ dà tutte le improprie ortocrone; $\Lambda_2\times\times SO^+(3,1)=SO^-(3,1)$, e infine $\Lambda_1\Lambda_2\times SO^+(3,1)$ dà tutte le improprie anticrone. Si pensa che solamente l'insieme $SO^+(3,1)$ possa corrispondere a osservatori fisicamente realizzabili. Ciò non toglie che lo studio delle simmetrie dei sistemi fisici rispetto alle trasformazioni Λ_1, Λ_2 abbia grandissimo interesse specialmente per quanto riguarda la teoria dei campi (si vedano gli articoli «Conservazione/invarianza», «Forza/campo», «Particella» in questa stessa *Enciclopedia*). È difficile immaginare due osservatori A e B collegati da Λ_2 in quanto il tempo in A scorrerebbe nella direzione opposta a quello in B. In modo simile, se Λ_1 collegasse A a B, un guanto destro visto da A apparirebbe sinistro se visto da B: la trasformazione Λ_1 opera quindi come la riflessione in uno specchio (simmetria di parità).

Sia poi $\mathbf{v}=|\mathbf{v}|\mathbf{n}$ una generica velocità, $\beta=|\mathbf{v}|/c$, $\gamma=1/\sqrt{1-\beta^2}$, $\beta^i=\beta n^i$, $i=1,2,3$. Si dirà *boost* lungo $\mathbf{v}$ la $\Lambda\in SO^+(3,1)$ data da $B(\mathbf{v})$:

$$(3.16)\qquad B^\alpha{}_\beta=\begin{Vmatrix} \gamma & \gamma\beta^1 & \gamma\beta^2 & \gamma\beta^3 \\ \gamma\beta^1 & 1+(n^1)^2(\gamma-1) & n^1n^2(\gamma-1) & n^1n^3(\gamma-1) \\ \gamma\beta^2 & n^1n^2(\gamma-1) & 1+(n^2)^2(\gamma-1) & n^2n^3(\gamma-1) \\ \gamma\beta^3 & n^1n^3(\gamma-1) & n^2n^3(\gamma-1) & 1+(n^3)^2(\gamma-1) \end{Vmatrix}.$$

Chiaramente, se $n^1=1$, $n^2=n^3=0$, si ritrova la (3.8), e se $n^1=n^3=0$, $n^2=1$ (oppure $n^1=n^2=0$, $n^3=1$) le (3.9). È facile verificare che $B\in SO^+(3,1)$ e che le $B(\mathbf{v})$ non formano un gruppo.

Si devono considerare come $\Lambda\in SO^+(3,1)$ le cosiddette «rotazioni», caratterizzate da $x'^0=x^0$, ossia $\Lambda^0{}_0=1$, $\Lambda^i{}_0=\Lambda^0{}_i=0$. In tal caso (cfr. (1.5), (1.6)),

$$(3.17)\qquad x'^i=\sum_{k=1}^{3}\Lambda^i{}_k x^k.$$

Denotando con R la sottomatrice di Λ con elementi $\Lambda^i{}_k$, $i,k=1,2,3$, si vede che (3.10) implica $R^TR=\mathbf{1}$ (matrice identità 3×3), ossia R è una matrice ortogonale in tre dimensioni. Queste matrici formano il gruppo $O(3)$ e dipendono da tre angoli (secondo Eulero) che si possono introdurre nel modo seguente. Siano

$$(3.18)\qquad R_3(\varphi)=\begin{vmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{vmatrix}\qquad R_2(\psi)=\begin{vmatrix} \cos\psi & 0 & \sin\psi \\ 0 & 1 & 0 \\ -\sin\psi & 0 & \cos\psi \end{vmatrix};$$

allora la generica R (con $\det R=1$) è data da $R(\vartheta,\varphi,\psi)=R_3(\varphi)\,R_2(\psi)\,R_3(\vartheta)$, dove φ, ψ, ϑ sono angoli generici in: $0\leq\varphi\leq 2\pi$, $0\leq\psi\leq\pi$, $0\leq\vartheta\leq 2\pi$. Fisicamente, una rotazione collega sistemi che differiscono solo per l'orientamento degli assi, ma che non hanno velocità relativa. È quindi logico attendersi che la piú generica $\Lambda\in SO^+(3,1)$ sia il prodotto di una rotazione per un *boost*.

3.2. Il calcolo tensoriale.

Nella discussione precedente è stata introdotta una serie di assunzioni:

Assunzione 1. I fenomeni fisici si svolgono nello spazio-tempo $\mathfrak{M}$.

Assunzione 2. Esistono in $\mathfrak{M}$ infiniti osservatori inerziali (OI): ad ogni OI, denotato con A, corrispondono coordinate x^α_A. Le coordinate x^α_A di A e x^α_B di un altro OI B sono collegate da una trasformazione di Poincaré

$$(3.19)\qquad x_A=\Lambda_{AB}x_B+\Xi_{AB},$$

con $\Lambda^T_{AB}g\Lambda_{AB}=g$, per cui $du^2=-dx^T_A\,g\,dx_A=-dx^T_B\,g\,dx_B$ è lo stesso per tutti gli OI. Questo fatto esprime la costanza della velocità della luce. Il principio di relatività richiede poi che:

Assunzione 3. Tutte le leggi della fisica si scrivono nello stesso modo per tutti gli OI. Il calcolo tensoriale è stato sviluppato per realizzare il principio di relatività e generalizza il calcolo vettoriale usato nella geometria e nella fisica in $\mathbf{R}^3$.

Assunzione 4. Le grandezze fisiche in un generico OI (ad esempio A) appaiono sempre raggruppate in insiemi della forma

$$(3.20)\qquad T^{\mu_1\ldots\mu_p}_{A,\,\nu_1\ldots\nu_q}(x_A)$$

dove $\mu_1\ldots\mu_p$, $\nu_1\ldots\nu_q$ sono indici sottoposti alle convenzioni 1), 2), 3) anticipate nel § 3.1. Piú specificamente, T è il simbolo e tutta l'espressione (3.20) è l'oggetto di cui in 1) del § 3.1.

Assunzione 5. Tra generici OI A e B si ha la relazione:

$$(3.21)\qquad T^{\mu_1\ldots\mu_p}_{A,\,\nu_1\ldots\nu_q}(x_A)=\Lambda^{\mu_1}_{AB,\,\alpha_1}\ldots\Lambda^{\mu_p}_{AB,\,\alpha_p}\Lambda^{\beta_1}_{BA,\,\nu_1}\ldots\Lambda^{\beta_q}_{BA,\,\nu_q}T^{\alpha_1\ldots\alpha_p}_{B,\,\beta_1\ldots\beta_q}(x_B)$$

dove x_A e x_B sono collegate dalla (3.19).

Osservazione 1: Si ha:

$$\Lambda^\mu_{AB,\,\alpha}=\frac{\partial x^\mu_A}{\partial x^\alpha_B}\qquad \Lambda^\beta_{BA,\,\nu}=\frac{\partial x^\beta_B}{\partial x^\nu_A}.$$

DEFINIZIONE 3. *Se $T^{(\mu)}_{A,(\nu)}(x_A)$ è dato in tutti gli OI e soddisfa alle (3.21), si dirà che esso è un tensore p volte controvariante e q volte covariante, o, brevemente, di tipo (p,q).*

Osservazione 2: Nella definizione 4 è stato abbreviato in (μ) l'insieme $\mu_1, \mu_2, \ldots, \mu_p$ e in (ν) l'insieme $\nu_1, \nu_2, \ldots, \nu_q$.

DEFINIZIONE 4. *Un tensore con un solo indice covariante (controvariante) dicesi vettore covariante (controvariante).*

DEFINIZIONE 5. *Un tensore senza indici dicesi scalare. Se T_A è uno scalare, allora $T_A(x_A) = T_B(x_B)$.*

Osservazione 3: Risulta chiaro che $T^{(\mu)}_{A,(\nu)}$ risulta assegnato tramite la (3.21) in tutti gli OI, se lo è in uno qualunque di essi.

Osservazione 4: Se $T^{(\mu)}_{(\nu)}$ ha componenti nulle per un OI, le ha nulle per un qualunque altro OI. Si dice che $T^{(\mu)}_{(\nu)} = 0$ è un tensore nullo. Eccone alcuni esempi.

Dalla relazione (3.10), riscritta nella forma $g_{\beta_1\beta_2}\Lambda^{\beta_1}_{BA,\nu_1}\Lambda^{\beta_2}_{BA,\nu_2} = g_{\nu_1\nu_2}$, risulta che $g_{A,\beta_1\beta_2}(x_A) = g_{\beta_1\beta_2}$ è un tensore, detto tensore metrico, due volte covariante. Si noti che $g_{A,\beta_1\beta_2}(x_A)$ è indipendente da A come conseguenza della definizione di $\Lambda \in O(3,1)$. Si ha inoltre

$$dx^\mu_A = \Lambda^\mu_{AB,\nu}\, dx^\nu_B \tag{3.22}$$

per cui dx^μ_A (ma non x^μ_A) è un vettore controvariante (tipo (1,0)).

Il simbolo di Kronecker $\delta^\mu_\nu = 1$ se $\mu = \nu$, $= 0$ se $\mu \neq \nu$, fornisce un altro esempio. Sia infatti $\delta^\mu_{A,\nu}(x_A) = \delta^\mu_\nu$; dalla (3.21) risulta: $\delta^\mu_{B,\nu}(x_B) = \delta^\mu_\nu$, per qualunque altro B.

Come ulteriore esempio è facile provare che $g^{\mu_1\mu_2}_A(x_A) = g^{\mu_1\mu_2}$ è un tensore di tipo (2,0). In quanto segue, se non sussistono ambiguità, si lascerà cadere l'indicazione dell'OI dal simbolo del tensore come pure l'indicazione esplicita della variabile x_A. La definizione di tensore è utile anche in quanto esistono regole per formare tensori a partire da altri tensori. Le regole sono:

REGOLA 1. Se $T^{(\mu)}_{(\nu)}$ e $S^{(\mu)}_{(\nu)}$ sono tensori dello stesso tipo, anche $R^{(\mu)}_{(\nu)} = \alpha T^{(\mu)}_{(\nu)} + \beta S^{(\mu)}_{(\nu)}$ è un tensore dello stesso tipo, dove α, β sono degli scalari; in termini piú espliciti, se $\alpha T^{\mu_1\ldots\mu_p}_{A,\nu_1\ldots\nu_q} + \beta S^{\mu_1\ldots\mu_p}_{A,\nu_1\ldots\nu_q}$ denota in A la componente $R^{\mu_1\ldots\mu_p}_{A,\nu_1\ldots\nu_q}$, in ogni altro OI si avrà $R^{\alpha_1\ldots\alpha_p}_{B,\beta_1\ldots\beta_q}$ univocamente determinato dalla (3.21) e uguale a $\alpha T^{\alpha_1\ldots\alpha_p}_{B,\beta_1\ldots\beta_q} + \beta S^{\alpha_1\ldots\alpha_p}_{B,\beta_1\ldots\beta_q}$.

REGOLA 2. Se $T^{(\mu)}_{(\nu)}$, $S^{(\mu)'}_{(\nu)}$ sono due tensori qualsiasi, e $(\mu) = (\mu_1, \ldots, \mu_p)$, $(\mu)' = (\mu_{p+1}\ldots\mu_t)$, $(\nu) = (\nu_1\ldots\nu_q)$, $(\nu)' = (\nu_{q+1}\ldots\nu_s)$, $T^{(\mu)}_{(\nu)}S^{(\mu)'}_{(\nu)'}$, è un tensore di tipo (t,s). Si tralascia la semplice dimostrazione. Ad esempio, se T^μ, S_ν sono vettori, allora $R^\mu_\nu = T^\mu S_\nu$ è un tensore di tipo (1,1).

REGOLA 3. Se $T^{\mu_1\ldots\mu_p}_{\nu_1\ldots\nu_q}$ è un tensore di tipo (p,q), allora $T^{\beta\mu_2\ldots\mu_p}_{\beta\nu_2\ldots\nu_q}$ è un tensore di tipo $(p-1, q-1)$.

Osservazione 5: Qui è implicita, secondo le convenzioni adottate, una sommatoria sull'indice muto β. Una espressione del tipo $T^{\mu_1\ldots\mu_p}_{\beta\beta\nu_3\ldots\nu_q}$ non rappresenta un tensore, ed è erronea, o priva di senso, o entrambe le cose insieme. Eccone alcuni esempi.

Sia $R^\mu_\nu = T^\mu S_\nu$; allora $T^\beta S_\beta$ è un tensore di tipo (0,0), ossia uno scalare,

cioè il prodotto scalare di T^μ e S_ν. Si noti che non si può formare direttamente il prodotto scalare di due vettori dello stesso tipo.

Per quanto detto, applicando successivamente le regole 2 e 3, la quantità $g_{\nu\beta}T^\beta = T_\nu$, di tipo (0,1), è costruita da T^β che è di tipo (1,0). Il tensore $g_{\nu\beta}$ serve dunque per *abbassare* gli indici e cambiare il tipo di tensore. L'operazione può ripetersi su tutti gli indici separatamente in un tensore qualsiasi. Inversamente, si ha $g^{\mu\alpha}T_\alpha = T^\mu$, per cui ci si riconduce al caso precedente. Invece l'effetto di δ^μ_ν è banale: $\delta^\mu_\beta T^\beta = T^\mu$, $\delta^\alpha_\nu T_\alpha = T_\nu$.

Osservazione 6: Secondo quanto visto in quest'ultimo esempio, valgono le formule tensoriali: $-du^2 = dx^\alpha\, dx_\alpha = dx^\alpha\, dx^\beta\, g_{\alpha\beta} = dx_\alpha\, dx_\beta\, g^{\alpha\beta}$.

Osservazione 7: Se T_ν, S_ν sono dello stesso tipo, è possibile definire prima $T^\mu = g^{\mu\nu}T_\nu$, e poi $T^\mu S_\mu = g^{\mu\nu}T_\nu S_\mu$ come prodotto scalare.

Le regole 1-3 varranno immutate anche nella relatività generale, mentre invece la regola seguente, che coinvolge differenziazioni, avrà necessità di alcune modifiche che risulteranno nella creazione del calcolo differenziale assoluto. Al momento, ci si limiterà alla formulazione:

REGOLA 4. Se $T^{\mu_1\ldots\mu_p}_{\nu_1\ldots\nu_q}(x)$ è un tensore, $T^{\mu_1\ldots\mu_p}_{\nu_1\ldots\nu_q\nu_{q+1}}(x) = \dfrac{\partial}{\partial x^{\nu_{q+1}}} T^{\mu_1\ldots\mu_p}_{\nu_1\ldots\nu_q}(x)$ è anche un tensore.

Osservazione 8: Nella regola 4, l'operatore $\partial/\partial x^\nu$ si comporta come un vettore covariante, che verrà scritto ∂_ν. Infatti, agendo su uno scalare T, si ha: $T_\nu = \partial_\nu T$, per cui T_ν rappresenta le componenti del gradiente di T. Ma dT, che è uno scalare, è dato da $dT = dx^\nu\, \partial_\nu T$, che esprime dT come il prodotto scalare del controvariante dx^ν e del covariante $\partial T/\partial x^\nu$.

REGOLA 5. Se $T^{(\mu)}_{(\nu)}$ è un tensore, anche $T^{(P\mu)}_{(P'\nu)}$ è un tensore, dove $(P\mu)$, $(P'\nu)$ sono permutazioni qualsiasi degli insiemi $(\mu) = \mu_1, \ldots, \mu_p$; $(\nu) = \nu_1, \ldots, \nu_q$.

Per esempio: se $T^{\mu_1\mu_2}$ è un tensore, anche $T^{\mu_2\mu_1}$ lo è; tuttavia, se T^μ_ν è un tensore, allora T^ν_μ ***non*** lo è piú.

Osservazione 9: Se si ha un tensore $T^{\mu_1\mu_2}$, facendo uso delle regole 1 e 5 si vede che anche $T^{\mu_1\mu_2} \pm T^{\mu_2\mu_1}$ sono tensori. Se $T^{\mu_1\mu_2} - T^{\mu_2\mu_1} = 0$, si dice che $T^{\mu_1\mu_2}$ è simmetrico, mentre se $T^{\mu_1\mu_2} = -T^{\mu_2\mu_1}$, si dice emisimmetrico. Queste proprietà vanno verificate per un solo OI, risultando poi valide in generale.

Per esempio: sia $\varepsilon_{\nu_1\nu_2\nu_3\nu_4}$ un tensore (generalmente noto come tensore di Levi-Civita) definito in A dalle seguenti condizioni: I) $\varepsilon_{0123} = 1$; II) $\varepsilon_{\nu_1\nu_2\nu_3\nu_4}$ è antisimmetrico nello scambio di due indici qualsiasi. ε è completamente definito dalle I) e II); inoltre esso rimane definito dalle I) e II) in tutti i sistemi B per i quali $\det\Lambda_{AB} = 1$. In generale si ha invece $\varepsilon_{A,0123} = (\det\Lambda_{AB})\varepsilon_{B,0123}$, come può essere facilmente verificato.

Il principio di relatività può essere ora formulato attraverso la seguente

Assunzione 6. Tutte le leggi fisiche si esprimono mediante la nullità di tensori costruiti dalle grandezze fisiche usando le regole 1-5 del calcolo tensoriale.

Cosí facendo, secondo l'osservazione 4, tutte le leggi vengono scritte esattamente nello stesso modo per tutti gli OI. Dovrebbe essere evidente che anche l'u-

guaglianza di tensori dello stesso tipo rientra nell'assunzione 6, in quanto essa implica la nullità della differenza di detti tensori, la quale, a norma della regola 1, è ancora un tensore dello stesso tipo.

Sorge quindi il problema d'identificare i tensori che rappresentano le grandezze fisiche. Sarà di un certo aiuto il considerare la formula $-du^2 = g_{\alpha\beta}\, dx^\alpha\, dx^\beta = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2$, come l'analogo in $\mathfrak{M}$ della distanza euclidea in $\mathbf{R}^3$: $ds^2 = dx^2 + dy^2 + dz^2$, con la sola, ma importante, differenza di segno davanti ai differenziali. Sotto questo punto di vista, le $\Lambda \in O(3,1)$ sono l'analogo in $\mathfrak{M}$ delle rotazioni $R \in O(3)$ in $\mathbf{R}^3$. La metrica du^2 in $\mathfrak{M}$ è dunque il punto di partenza per uno studio geometrico in $\mathfrak{M}$, analogo alla geometria solida in $\mathbf{R}^3$, in cui le velocità $\mathbf{v}$ in $B(\mathbf{v})$ svolgono il ruolo di angoli iperbolici.

4. *Dinamica relativistica ed equazioni di Maxwell.*

4.1. Tempo proprio.

Nella relatività di Einstein il moto di una particella P viene ancora descritto assegnando la posizione di P, $x^i(t)$, in funzione del tempo. Tuttavia la velocità

$$\mathbf{v} = \frac{d\mathbf{x}}{dt} \tag{4.1}$$

non ha alcun riscontro in una grandezza tensoriale. Infatti la legge di trasformazione delle velocità non è altro che il teorema di addizione della velocità $\mathbf{v}$ con la velocità relativa dei sistemi A e B (2.23). Questa legge risulta altamente nonlineare. Inoltre dt è la componente dx^0/c di un vettore e il rapporto tra componenti diverse di un vettore non rientra nelle regole di calcolo tensoriale esposte nel § 3. Per queste ragioni si usa invece la quadrivelocità

$$u^\mu = c\frac{dx^\mu}{du} \tag{4.2}$$

dove

$$-du^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu. \tag{4.3}$$

È facile controllare che, poiché $du = c\, dt \sqrt{1-\beta^2}$,

$$u^0 = \frac{c}{\sqrt{1-\beta^2}} \qquad u^i = \frac{v^i}{\sqrt{1-\beta^2}}. \tag{4.4}$$

Inoltre

$$g_{\mu\nu} u^\mu u^\nu = -c^2. \tag{4.5}$$

In tutte queste formule s'intende $|\beta| < 1$, altrimenti esse perderebbero di significato. Questa condizione non dipende dal sistema di riferimento. Essa implica infatti $du^2 > 0$, dato che

$$du^2 = (dx^0)^2 - |d\mathbf{x}|^2 = dt^2\left(c^2 - \left|\frac{d\mathbf{x}}{dt}\right|^2\right) = dt^2(1-\beta^2)c^2.$$

Ma du^2 è invariante e cosí pure $|\beta| < 1$. Si supponga che P proceda di moto uniforme. Ponendosi nel sistema a riposo di P, si ottiene $d\tau \equiv (1/c)\, du = dt$, per cui $d\tau$ misura lo scorrere del tempo nel sistema di P, detto anche tempo proprio. Ne segue anche che $d\tau = dt\sqrt{1-\beta^2} = du/c$ per ogni corpo in moto con velocità uniforme $v = \beta c$, e che non occorre dunque porsi nel sistema di P per calcolarlo. Per corpi non sottoposti ad accelerazioni troppo intense, si può assumere come tempo proprio ΔT decorso fra due eventi E_1, E_2 la quantità

$$\Delta T = \int_{E_1}^{E_2} dt \sqrt{1-\beta^2} \tag{4.6}$$

che è sempre minore della differenza $t_2 - t_1$ fra i tempi di E_1, E_2 per via del fattore $\sqrt{1-\beta^2} < 1$ che deriva dalla dilatazione dei tempi.

Dato un evento $\mathcal{O} \in \mathfrak{M}$, di coordinate ad esempio $x^\alpha_{\mathcal{O}} = 0$, tutti gli eventi di $\mathfrak{M}$ possono essere classificati nel modo seguente in cinque sottoinsiemi:

1) $\mathfrak{M}^+_{\mathcal{O}}$, il *cono futuro* di $\mathcal{O}$. È costituito dagli eventi raggiungibili da un segnale partito da $\mathcal{O}$ e viaggiante con velocità $\mathbf{v}$, dove $|\mathbf{v}| < c$; per essi: $x^0 > 0$, $-du^2 = g_{\mu\nu} x^\mu x^\nu \leq 0$;
2) $\mathfrak{M}^-_{\mathcal{O}}$, il *cono passato* di $\mathcal{O}$: $x^0 < 0$, $du^2 \geq 0$;
3) $\mathfrak{M}^0_{\mathcal{O}}$, il presente di $\mathcal{O}$: $\mathfrak{M}^0_{\mathcal{O}} = \mathfrak{M} - (\mathfrak{M}^+_{\mathcal{O}} \cup \mathfrak{M}^-_{\mathcal{O}})$;
4) $\partial\mathfrak{M}^+_{\mathcal{O}}$, il *cono luce futuro* di $\mathcal{O}$: $x^0 > 0$, $du^2 = 0$;
5) $\partial\mathfrak{M}^-_{\mathcal{O}}$, il *cono luce passato* di $\mathcal{O}$: $x^0 < 0$, $du^2 = 0$.

Nel piano x^0, x^1, il cono $\partial\mathfrak{M}^+_{\mathcal{O}} \cup \partial\mathfrak{M}^-_{\mathcal{O}}$ si riduce alle bisettrici degli assi: $x^0 = \pm x^1$. Gli eventi in $\mathfrak{M}^+_{\mathcal{O}}(\mathfrak{M}^-_{\mathcal{O}})$ sono nel futuro (passato) di $\mathcal{O}$ per qualunque oi. Quelli in $\mathfrak{M}^0_{\mathcal{O}}$ invece possono essere visti nel passato o nel futuro di $\mathcal{O}$ a seconda dell'oi. Se $A \in \mathfrak{M}^\pm_{\mathcal{O}}$, si dice che A, $\mathcal{O}$ sono separati da un intervallo di tipo tempo, se $A \in \mathfrak{M}^0_{\mathcal{O}}$ di tipo spazio, se $A \in \partial\mathfrak{M}^\pm_{\mathcal{O}}$ di tipo nullo o luce.

4.2. Energia e impulso.

Occorre ora giungere a una definizione soddisfacente di energia E e di quantità di moto $\mathbf{p}$ di un punto materiale libero P. Nella relatività galileiana

$$\mathbf{p} = m\mathbf{v} \qquad E = mv^2/2 = T, \tag{4.7}$$

dove m è la massa di P e T l'energia cinetica. Le stesse obiezioni già discusse per $\mathbf{v}$ all'inizio del § 4.1, valgono per $\mathbf{p}$ ed E, che non paiono dunque rivestire ovvie proprietà tensoriali nel caso einsteiniano. Per esse si prospetta dunque una modificazione anche e soprattutto in vista del fatto che le $\mathbf{p}$, E sono quantità conservate per un sistema isolato e sono centrali nella formulazione hamiltoniana delle equazioni del moto. Per queste ragioni il loro ruolo relativistico va discusso in modo esauriente.

Per definire in modo corretto $\mathbf{p}$, E, occorre rifarsi alla loro definizione in base al formalismo canonico [Goldstein 1959]. Si assumerà quindi l'esistenza di una *lagrangiana* $\mathcal{L}(\mathbf{x}, d\mathbf{x}/dt)$ tale che il principio di minima azione

$$(4.8)\qquad \delta\int_1^2 \mathcal{L}(\mathbf{x},\mathbf{v})\,dt=0$$

tra due eventi 1, 2, determini le equazioni del moto

$$(4.9)\qquad \frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial v^i}\right)-\frac{\partial\mathcal{L}}{\partial x^i}=0.$$

Seguendo il procedimento solito, si possono introdurre il momento

$$(4.10)\qquad p^i=\frac{\partial\mathcal{L}}{\partial v^i}$$

e la hamiltoniana

$$(4.11)\qquad H=p^iv^i-\mathcal{L}.$$

Nel caso di una particella libera si ammetterà che $\mathcal{L}$ sia indipendente da $\mathbf{x}$ e che sia invariante almeno rispetto alle rotazioni; pertanto $\mathcal{L}=f(v^2)$, e dunque: $p^i=2f'v^i$, $H=2f'(v^i)^2-f$. Si richiederà inoltre che l'azione tra gli eventi 1, 2, non dipenda dall'01; in tal modo pure le equazioni del moto dedotte dalle (4.8) non dipenderanno dalla scelta dell'01 e sarà stato realizzato il principio di relatività.

Dunque, l'azione $S=\int_1^2 f(v^2)\,dt$ non deve dipendere dall'01; si può anche scrivere

$$S=\int_1^2 du\frac{f(v^2)}{c\sqrt{1-\beta^2}}$$

dove, per quanto detto, du è invariante. Ne segue che $f(v^2)/c\sqrt{1-\beta^2}$ deve essere pure invariante. Si può poi sempre portare P a riposo in un punto dato della traiettoria, ottenendo $v=0$, per cui

$$\frac{f(v^2)}{c\sqrt{1-\beta^2}}=\frac{f(0)}{c}=\text{cost.}$$

Ne segue

$$S=\int_1^2 du\frac{f(0)}{c}=\int_1^2 dt\sqrt{1-\beta^2}\,f(0),$$

da cui

$$f(v^2)=f(0)\sqrt{1-\beta^2}\qquad p^i=-\frac{f(0)}{c^2}\frac{v^i}{\sqrt{1-\beta^2}}\qquad H=-\frac{f(0)}{\sqrt{1-\beta^2}}.$$

Ma per piccole velocità, la nuova definizione deve coincidere con quella galileiana; pertanto, se $|\beta|\ll 1$,

$$p^i\sim-\frac{f(0)}{c^2}v^i\sim mv^i\qquad f(0)=-mc^2.$$

Si ricava allora

$$(4.12)\qquad H=\frac{mc^2}{\sqrt{1-\beta^2}}=\sqrt{p^2c^2+m^2c^4}=E(p)\qquad p^i=\frac{mv^i}{\sqrt{1-\beta^2}},$$

che sembra a prima vista molto diversa dalla (4.7). Tuttavia E è determinata solo a meno di una costante, e si vede infatti che $H|_{v=0}=mc^2$, per cui l'energia cinetica è

$$(4.13)\qquad T=\frac{mc^2}{\sqrt{1-\beta^2}}-mc^2=E(p)-mc^2.$$

Sviluppando (4.13) in serie di potenze di β^2 si trova infatti

$$T=\frac{mv^2}{2}+\frac{3}{8}\frac{mv^4}{c^2}+\ldots,$$

il cui termine dominante è in accordo con (4.7).

T definita tramite la (4.13) non è di grande utilità; piuttosto, è facile vedere che $p^i=mu^i$, $H=T+mc^2=mcu^0$. Ponendo allora $p^0=mu^0=H/c$, risulta che p^α forma un quadrivettore parallelo a u^α. Inoltre $-g_{\alpha\beta}p^\alpha p^\beta=m^2c^2$, ossia

$$(4.14)\qquad (p^0)^2-|\mathbf{p}|^2=m^2c^2\qquad p^0=\sqrt{p^2+m^2c^2}.$$

Nella meccanica newtoniana si ha la famosa relazione

$$(4.15)\qquad \mathbf{F}=m\mathbf{a}=d\mathbf{p}/dt$$

tra forza ed accelerazione. Dalla (4.15), moltiplicando per $\mathbf{v}$, si ottiene $\mathbf{F}\cdot\mathbf{v}=\mathbf{v}\cdot(d\mathbf{p}/dt)$; il termine a sinistra è il lavoro per unità di tempo delle forze che agiscono su P e deve uguagliare il cambiamento di energia cinetica, cioè $dT/dt=\mathbf{v}\cdot(d\mathbf{p}/dt)$, che è naturalmente in accordo con la (4.7). Ma è in accordo anche con la definizione relativistica. Infatti, ponendo $p^0=(T+mc^2)/c$, si ottiene $c(dp^0/dt)=v^i(dp^i/dt)$ e moltiplicando per $\gamma\equiv 1/\sqrt{1-\beta^2}$: $u^0(dp^0/dt)-u^i(dp^i/dt)=0$; dividendo ora per m: $u^0(du^0/dt)-u^i(du^i/dt)=0$, ossia

$$\frac{1}{2}\frac{d}{dt}[(u^0)^2-(\mathbf{u})^2]=\frac{1}{2}\frac{d}{dt}c^2\equiv 0,$$

che risulta verificata identicamente.

Osservazione 1: Talvolta si continua a scrivere $\mathbf{p}=m\mathbf{v}$, dove $m=m_0/\sqrt{1-\beta^2}$ è la massa in movimento, distinta dalla massa a riposo m_0. Quando $\beta\to 1$, la massa m tende a infinito e cosí pure l'inerzia del corpo P. Risulta dunque sempre piú difficile accelerare corpi quando la loro velocità si avvicina a c e l'incremento d'inerzia è tale da vanificare ogni tentativo di superare c.

Osservazione 2: La relazione $p^\alpha p_\alpha=-m^2c^2$ per una particella libera P viene anche chiamata gergalmente il *mass-shell* per P.

Osservazione 3: Il fotone o quanto di luce richiede una trattazione separata in quanto per esso $\beta=1$ e quindi la $p^\alpha=mu^\alpha$ non si applica. Se tuttavia in $p^i=mv^i/\sqrt{1-\beta^2}$ s'immaginano $\beta\to 1$ e $m\to 0$ in modo tale da mantenere p^i finito, si arriva al fotone come caso limite; ricordando la (4.14), al limite si ha $p^0=|\mathbf{p}|$. La velocità $\mathbf{v}$ del fotone, dato che $|\mathbf{v}|=c$, non serve che a specificare la

direzione di $\mathbf{p}$ ma non è legata a $|\mathbf{p}|$. La relazione $p^0=|\mathbf{p}|$ e la $E=c|\mathbf{p}|$ che ne segue sono anche deducibili dalla relazione di De Broglie

$$p=\frac{h}{\lambda} \tag{4.16}$$

e dalla relazione di Planck

$$E=h\nu \tag{4.17}$$

dove λ e ν sono la lunghezza d'onda e la frequenza dell'onda associata al fotone; infatti $\nu\lambda=c$.

4.3. Equivalenza massa-energia.

Fintantoché si tratta con una particella singola o con processi dove non vengono create o distrutte particelle e non esistono processi eso- o endotermici (cioè non esistono processi inelastici), la somma delle energie cinetiche iniziali uguaglia la somma di quelle finali, e cosí pure la somma delle p^0 iniziali è uguale alla somma di quelle finali, dato che queste due somme differiscono solo per il contributo mc^2 di ciascuna particella che, per quanto detto, non varia. Se invece entrano in gioco processi inelastici, oppure creazione o distruzione di particelle, occorre procedere con estrema cautela, in quanto i risultati galileiani non servono piú da guida sicura.

Della fisica galileiana, o meglio della meccanica di Newton, verrà tuttavia mantenuto il principio di azione e reazione, che appare verificato sperimentalmente con grande accuratezza. Per due corpi P_1 e P_2 con momenti $\mathbf{p}_1$, $\mathbf{p}_2$, le equazioni del moto sono

$$\frac{d\mathbf{p}_1}{dt}=\mathbf{F}_{12} \qquad \frac{d\mathbf{p}_2}{dt}=\mathbf{F}_{21}$$

dove $\mathbf{F}_{12}(\mathbf{F}_{21})$ è la forza con cui il corpo $P_2(P_1)$ agisce su $P_1(P_2)$. Secondo il principio di azione e reazione si ha sempre $\mathbf{F}_{12}+\mathbf{F}_{21}=0$, qualunque sia la natura delle forze; pertanto $d(\mathbf{p}_1+\mathbf{p}_2)/dt=0$, ossia $\mathbf{p}_1+\mathbf{p}_2$ è una costante del moto.

Generalizzando questo concetto, si può assumere che in una collisione di n corpi iniziali, etichettati dall'indice i, dalla quale emergono m corpi etichettati dall'indice f, valga la legge di conservazione:

$$\Delta p^k=\sum_i p_i^k-\sum_f p_f^k=0 \qquad k=1,2,3. \tag{4.18}$$

Assumendo che la (4.18) sia valida per tutti gli 01, si deve avere anche $\Delta p^0=0$ per tutti gli 01; infatti, usando per i singoli corpi trasformazioni di Lorentz del tipo (2.14) e imponendo che sia $\Delta p^1=\Delta p'^1=0$, si ottiene $\Delta p^0=\Delta p'^0=0$. Va quindi assunta la validità di

$$\sum_i p_i^0=\sum_f p_f^0 \tag{4.19}$$

per qualunque processo isolato, nel quale, cioè, non avviene scambio d'impulso ed energia con un campo esterno.

La (4.19) significa ben di piú della solita conservazione galileiana. Infatti la si può riscrivere nella forma

$$\begin{aligned}\text{energia cinetica apparsa}=\Delta E&=\sum_f T_f-\sum_i T_i=\Big(\sum_i m_i-\sum_f m_f\Big)c^2=\\&=\Delta Mc^2=\text{massa sparita}\times c^2.\end{aligned} \tag{4.20}$$

ΔMc^2 ha la dimensione di un'energia e viene chiamato l'equivalente energetico di ΔM. Un grammo ha come equivalente circa $9\cdot 10^{13}$ joule, ovvero circa $2{,}5\cdot 10^7$ Kwh, una quantità enorme di energia. Prendendo letteralmente la (4.20), la sparizione ΔM di materia corrisponde all'aumento di energia $\Delta E=\Delta Mc^2$.

La (4.20) non trova posto nella meccanica galileiana poiché nella fisica e nella chimica prerelativistiche si ha sempre assoluta conservazione della massa. Nelle reazioni chimiche ordinarie non si ha mai distruzione o creazione di particelle massive, ma solamente l'aggregarsi o il rompersi di molecole piú o meno complesse e l'emissione di fotoni. Le variazioni sono di fatto empiricamente nulle entro gli errori sperimentali che difficilmente scendono (o comunque scendevano agli esordi della relatività) al di sotto di una parte in 10^8. Ora, se tutta l'energia chimica provenisse da ΔMc^2, sarebbe facile convincersi che i ΔM corrispondenti alle piú violente reazioni chimiche conosciute sono troppo piccoli per essere misurati. Ad esempio, se 12 grammi di carbonio reagiscono con 32 grammi di ossigeno per formare 44 grammi di CO_2, si sviluppano $94{,}03\cdot 10^3$ calorie, ossia circa $4\cdot 10^5$ joule. L'equivalente in massa è dato da $\Delta M\sim 4\cdot 10^{-9}$ g, per cui $\Delta M/M\sim 10^{-10}$, ben al di sotto delle bilance piú sensibili. Per questa ragione, anche assumendo che l'energia chimica altro non sia che ΔMc^2 e quindi assumendo sparizione di massa, non si entra in conflitto con l'evidenza derivata dalle reazioni chimiche, che hanno rappresentato la quasi totalità delle reazioni osservabili all'avvento della relatività.

La scoperta delle reazioni nucleari ha cambiato completamente la situazione. Nella fissione dell'uranio 235 sparisce una massa corrispondente a circa un millesimo del totale; nelle complesse catene di reazioni termonucleari che avvengono nell'interno delle stelle, viene convertito idrogeno in elio. Il peso atomico dell'idrogeno è 1,008, per cui quattro protoni hanno un peso atomico totale 4,032, mentre l'elio pesa solo 4,004. Fondendo quattro protoni in una particella, si perde quindi $\Delta M/M=0{,}028/4=0{,}7$ per cento, che si ritrova in energia. Si giunge a casi estremi considerando l'annichilazione di materia o antimateria. Ad esempio, un antiprotone $\bar{P}$ può annichilare un protone P producendo solamente fotoni: $P+\bar{P}=n\gamma$. Nello stato finale tutto è energia cinetica e quindi energia utilizzabile; in questo caso, il 100 per cento della massa iniziale è sparita. Purtroppo, per fabbricare un antiprotone con le inefficientissime macchine acceleratrici moderne si spende molta piú energia (per un fattore di miliardi) di quella che si spera di ricavare dalla sua annichilazione.

Si assumerà quindi la (4.19) come forma corretta della conservazione dell'energia, che è possibile porre in forma tensoriale

$$\sum_i p_i^\alpha=\sum_f p_f^\alpha \tag{4.21}$$

valida per ogni OI. Si riguarderanno tutte le leggi prerelativistiche di conservazione della massa e il concetto di energia chimica come un'approssimazione della (4.21), valida quando $\Delta M/M \ll 1$.

4.4. Aberrazione della luce ed effetto Doppler.

Sia p l'impulso proveniente da una stella. Per semplicità si supporrà $p^1 = -p^0 \cos\vartheta$, $p^2 = -p^0 \sin\vartheta$, $p^3 = 0$ in un certo riferimento. Ci si ponga ora in un OI in moto rispetto al precedente lungo l'asse x con velocità $v = \beta c$. Si avrà

$$p'^0 = \frac{p^0 - \beta p^1}{\sqrt{1-\beta^2}} = +p^0 \frac{1+\beta\cos\vartheta}{\sqrt{1-\beta^2}} \qquad p'^1 = \frac{p^1 - \beta p^0}{\sqrt{1-\beta^2}} = -p^0 \frac{\cos\vartheta + \beta}{\sqrt{1-\beta^2}}$$

$$p'^2 = p^2 = -p^0 \sin\vartheta,$$

per cui

$$\cos\vartheta' = -\frac{p'^1}{p'^0} = \frac{\cos\vartheta + \beta}{1+\beta\cos\vartheta}, \tag{4.22}$$

cioè la luce della stella appare provenire da una direzione caratterizzata da ϑ'. Se $\beta \gtrless 0$, si ha $\cos\vartheta' \gtrless \cos\vartheta$ e quindi $\vartheta' \lessgtr \vartheta$. Per $\beta > 0$ e prossimo a 1, $\cos\vartheta'$ è prossimo a 1 e ϑ' è prossimo a 0; questo significa che, aumentando la velocità v nella direzione positiva di x, le immagini delle stelle paiono addensarsi lungo la stessa direzione. Il fenomeno ha una certa somiglianza con quello osservato durante una fitta pioggia; per chi corre, le gocce di pioggia sembrano venire incontro e provenire dalla direzione anteriore.

Durante una rivoluzione intorno al Sole, la Terra cambia continuamente velocità, per cui le immagini delle stelle sembrano provenire da direzioni diverse durante l'anno e descrivono delle ellissi apparenti sulla volta celeste.

Le stesse formule (4.22) si prestano per valutare l'effetto Doppler. Questo consiste nel cambiamento della frequenza di un'onda elettromagnetica quando la sorgente è in moto rispetto all'osservatore. Si ha quindi, tenendo conto che $p^0 = E/c = h\nu/c$, secondo quanto visto

$$\nu' = \nu \frac{1+\beta\cos\vartheta}{\sqrt{1-\beta^2}}. \tag{4.23}$$

Da questa relazione si vede che, se $\cos\vartheta = 1$, e cioè se si va incontro alla sorgente, $\nu' = \nu\sqrt{1+\beta}/\sqrt{1-\beta}$, ossia $\nu' > \nu$ se $\beta < 0$; se ci si allontana: $\nu' < \nu$, $\beta > 0$. Ma anche se $\cos\vartheta' = 0$, si ha $\cos\vartheta = -\beta$ da (4.22) e $\nu' = \nu\sqrt{1-\beta^2} < \nu$, per cui l'effetto persiste anche se la velocità è normale alla direzione di arrivo dei fotoni. In generale

$$\nu' = \frac{\sqrt{1-\beta^2}}{1-\beta\cos\vartheta'}\nu. \tag{4.24}$$

L'effetto è presente anche nelle onde sonore. Il fischio di una locomotiva in arrivo è piú acuto di quello di una in allontanamento. La formula è tuttavia

diversa. La sorgente emette infatti una cresta d'onda ogni intervallo T (periodo) muovendosi verso l'osservatore con velocità v. Sia V la velocità del suono. Durante il tempo T la sorgente si sposta in avanti di Tv, per cui, verso l'osservatore, la distanza fra due creste successive è $TV - Tv = \lambda'$, dove λ' è la lunghezza d'onda percepita. Dunque: $\lambda' = \lambda(1 - v/V) = \lambda(1-\beta)$. Nel caso della luce il periodo della sorgente appare dilatato in $T' = T/\sqrt{1-\beta^2}$, per cui $\lambda' = T'(c-v) = \lambda\sqrt{1-\beta}/\sqrt{1+\beta}$, come risulta dalla (4.24). Il cambiamento è necessario per il principio di relatività. Scambiando λ, ν, ϑ, β con λ', ν', ϑ', $-\beta$, tutte le formule proposte rimangono invariate per la simmetria fra i due OI. Nel caso delle onde acustiche, l'OI fermo rispetto all'aria è privilegiato e non c'è simmetria.

4.5. Effetto Fizeau.

In un mezzo con indice di rifrazione n la luce si propaga con velocità c/n. Supponendo di porre il mezzo in moto con velocità v, si otterrà la velocità V della luce rispetto al laboratorio componendo queste due velocità secondo la (2.23). Il risultato è

$$V = (c/n + v)/(1 + v/cn) \approx \frac{c}{n} + v(1 - 1/n^2), \tag{4.25}$$

che differisce da quello non relativistico dato da $V = c/n + v$. Ciò veniva interpretato come un parziale trascinamento dell'etere da parte del mezzo in movimento secondo schemi piú o meno artificiosi. In realtà si trattava di una incorretta applicazione del teorema di composizione delle velocità. Se n dipende dalla lunghezza d'onda λ, l'effetto Doppler cambia λ secondo quanto detto precedentemente e questo introduce un termine correttivo nella (4.25), già predetto da Lorentz ed osservato da Zeeman. Il risultato è

$$V = \frac{c}{n} + v\left(1 - \frac{1}{n^2} - \frac{\lambda}{n}\frac{dn}{d\lambda}\right). \tag{4.26}$$

4.6. Elettrodinamica classica.

Verranno ora richiamati alcune definizioni e risultati standard dell'elettrodinamica classica nel vuoto. Secondo Maxwell il campo elettromagnetico (CEM) è descritto da due vettori, **E** (campo elettrico) e **B** (campo magnetico), che obbediscono, secondo quanto già detto nel § 1, alle seguenti equazioni (scritte in unità gaussiane):

$$\operatorname{div}\mathbf{B} = 0 \tag{4.27.1}$$

$$\operatorname{rot}\mathbf{E} = -\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t} \tag{4.27.2}$$

dette omogenee, e alle

$$\text{(4.27.3)} \qquad \operatorname{div}\mathbf{E}=4\pi\rho$$

$$\text{(4.27.4)} \qquad \operatorname{rot}\mathbf{B}=\frac{1}{c}\frac{\partial\mathbf{E}}{\partial t}+\frac{4\pi}{c}\mathbf{j}$$

dette equazioni di Maxwell non omogenee. In queste ultime appaiono le funzioni $\rho(\mathbf{x})$ (densità di carica elettrica) e $\mathbf{j}$ (densità di corrente elettrica) che descrivono la distribuzione delle cariche nello spazio. Prima di discutere le proprietà di $\mathbf{E}$, $\mathbf{B}$, è piú conveniente esaminare ρ e $\mathbf{j}$.

Per mostrare in qual modo ρ, $\mathbf{j}$ sono legate al moto delle cariche, si pensa a una carica come a un corpo quasi puntiforme, dotato, oltreché di massa m (di cui ci si disinteresserà fino al momento di occuparsi delle equazioni del moto), anche di un attributo supplementare: la carica Q.

Sia data da $\xi^\alpha(\tau)$ (oppure da $\xi^\alpha(t)$ se si sceglie $t\equiv\tau$) la traiettoria Γ della carica nello spazio-tempo in funzione di un parametro monotòno lungo la traiettoria, e si supponga per semplicità che in quiete la carica abbia forma di una sferetta di raggio R. La densità di carica nell'interno sarà Q/Ω, dove $\Omega=4\pi R^3/3$ è il volume della sferetta. Al limite $R\to 0$, la densità sarà data da una funzione quasi dappertutto nulla, ma che diventa infinita sulla carica. Si indicherà con $Q\delta^{(3)}(\mathbf{x})$ una tale funzione se la carica si trova nell'origine, dove

$$\text{(4.28)} \qquad \int_{\mathbf{R}^3}\delta^{(3)}(\mathbf{x})\,d^3\mathbf{x}=1.$$

La $\delta^{(3)}(\mathbf{x})$ è stata introdotta per la prima volta da Dirac [1930]. Essa non è in realtà una funzione, bensí appartiene alla classe piú ampia delle distribuzioni. Se la carica si trova nel punto $\boldsymbol{\xi}$, si scriverà $\rho(\mathbf{x})=Q\delta^{(3)}(\mathbf{x}-\boldsymbol{\xi})$.

La funzione di Dirac δ può definirsi su di uno spazio $\mathbf{R}^n$ qualunque tramite l'identità

$$\text{(4.29)} \qquad f(x)=\int_{\mathbf{R}^n}\delta^{(n)}(x-\xi)f(\xi)\,d^n\xi \qquad x,\,\xi\in\mathbf{R}^n,$$

che generalizza la (4.28). Se $\delta(x)$ è la funzione di Dirac in una dimensione, si ha

$$\text{(4.30)} \qquad \delta^{(n)}(x-\xi)=\prod_{i=1}^{n}\delta(x^i-\xi^i).$$

Accanto alla $\rho(\mathbf{x})$, s'introdurrà, per una carica in moto con velocità $\mathbf{v}$, il vettore corrente

$$\text{(4.31)} \qquad \mathbf{j}(\mathbf{x})=Q\mathbf{v}\delta^{(3)}(\mathbf{x}-\boldsymbol{\xi}(t)) \qquad \rho(\mathbf{x})=Q\delta^{(3)}(\mathbf{x}-\boldsymbol{\xi}(t))$$

con $\boldsymbol{\xi}(t)=\mathbf{v}t+\boldsymbol{\eta}$, $\boldsymbol{\eta}$ costante. Si ha perciò l'equazione di continuità

$$\text{(4.32)} \qquad \frac{\partial\rho}{\partial t}+\operatorname{div}\mathbf{j}=0,$$

avendo usato la formula di derivazione composta e $d\xi^i/dt=v^i$.

Se la carica Q è in moto *non* uniforme, è possibile generalizzare le (4.31) e (4.32) in modo tale da evidenziare anche l'invarianza relativistica della defini-

zione. S'introduca all'uopo la funzione $\delta^{(4)}(x-\xi)$ in $\mathfrak{M}$, dove $x,\xi\in\mathfrak{M}$, e sia data da $\xi^\alpha(\tau)$ l'evoluzione di Q in $\mathfrak{M}$. Considerando gli integrali

$$\text{(4.33)} \qquad j^\alpha(x)=cQ\int_\Gamma\delta^{(4)}(x-\xi(\tau))\frac{d\xi^\alpha}{d\tau}\,d\tau=cQ\int_\Gamma\delta^{(4)}(x-\xi(\tau))\,d\xi^\alpha,$$

e supponendo Q in moto uniforme ($\xi^0=\tau c$, $\xi^i=v^i\tau+\eta^i$), se si usa ripetutamente la (4.29) nella sola variabile τ si ottiene

$$j^0(x)=cQ\int_\Gamma\delta^{(3)}(\mathbf{x}-\mathbf{v}\tau-\boldsymbol{\eta})\delta(x^0-c\tau)c\,d\tau=cQ\delta^{(3)}\left(\mathbf{x}-\frac{\mathbf{v}}{c}x^0-\boldsymbol{\eta}\right),$$

per cui $j^0(x)/c$ coincide con $\rho(x)$. Inoltre $j^i(x)$ concorda con la definizione di $\mathbf{j}$:

$$j^i(x)=cQ\int_\Gamma\delta^{(3)}(\mathbf{x}-\mathbf{v}\tau-\boldsymbol{\eta})\delta(x^0-c\tau)v^i\,d\tau=Q\delta^{(3)}(\mathbf{x}-\boldsymbol{\xi}(x^0))\mathbf{v}.$$

L'equazione di continuità segue dalla (4.33) in quanto

$$\frac{\partial\rho}{\partial t}+\operatorname{div}\mathbf{j}=\frac{\partial j^\alpha}{\partial x^\alpha}=-cQ\int_\Gamma\frac{d}{d\tau}\delta^{(4)}(x-\xi(\tau))\,d\tau=-cQ\delta^{(4)}(x-\xi(\tau))\Big|_{\tau\to-\infty}^{\tau\to+\infty}=0,$$

poiché, essendo $\xi^0(\tau)=c\tau$, la funzione $\delta(x^0-c\tau)$ contenuta in $\delta^{(4)}$ è certamente nulla al limite. Si adotterà dunque la (4.33) come definizione di corrente anche quando il moto non è uniforme. Dall'analisi precedente è quasi ovvio constatare che $j^\alpha(x)$ è un vettore, e che l'equazione di continuità

$$\text{(4.34)} \qquad \frac{\partial j^\alpha}{\partial x^\alpha}=0$$

ha significato invariante.

Occorre ora legare questo risultato alla covarianza delle equazioni di Maxwell; si potrebbe dare il risultato, e poi verificarlo: invece è piú istruttivo seguire un procedimento deduttivo.

Anzitutto si constata che, se $\operatorname{div}\mathbf{B}=0$, deve esistere almeno un vettore $\mathbf{A}$ (vettore potenziale) tale che $\mathbf{B}=\operatorname{rot}\mathbf{A}$, ed uno scalare Φ tale che

$$\text{(4.35)} \qquad \mathbf{E}=-\frac{1}{c}\frac{\partial\mathbf{A}}{\partial t}-\operatorname{grad}\Phi.$$

$\mathbf{A}$ e Φ non sono determinati univocamente da (4.27.1) e (4.27.2); questa libertà addizionale può essere impiegata per imporre che sia soddisfatta la condizione di Lorentz

$$\text{(4.36)} \qquad \frac{1}{c}\frac{\partial\Phi}{\partial t}+\operatorname{div}\mathbf{A}=0.$$

Introducendo l'operatore d'onda

$$\text{(4.37)} \qquad \Box=g_{\alpha\beta}\frac{\partial^2}{\partial x^\alpha\,\partial x^\beta}$$

che è relativisticamente invariante, e la notazione A^α, con $A^0 = \Phi$, dalle precedenti relazioni si vede che (4.27.3) e (4.27.4) diventano

$$\Box A^\alpha = -\frac{4\pi}{c} j^\alpha; \tag{4.38}$$

pertanto A^α è un quadrivettore.

Inoltre, introducendo il tensore del campo elettromagnetico

$$F_{\alpha\beta} = \frac{\partial A_\beta}{\partial x^\alpha} - \frac{\partial A_\alpha}{\partial x^\beta} = -F_{\beta\alpha} \tag{4.39}$$

di componenti

$$F^{\alpha\beta} = \begin{vmatrix} 0 & E^1 & E^2 & E^3 \\ -E^1 & 0 & B^3 & -B^2 \\ -E^2 & -B^3 & 0 & B^1 \\ -E^3 & B^2 & -B^1 & 0 \end{vmatrix} \tag{4.40}$$

si vede che le (4.27.1) e (4.27.2) si scrivono

$$\varepsilon^{\lambda\alpha\beta\gamma} \frac{\partial F_{\alpha\beta}}{\partial x^\gamma} = 0 \tag{4.41}$$

e le (4.27.3) e (4.27.4)

$$\frac{\partial F^{\alpha\beta}}{\partial x^\beta} = 4\pi \frac{j^\alpha}{c}. \tag{4.42}$$

Invece dei due vettori **E**, **B** in $\mathbf{R}^3$, le equazioni di Maxwell vengono cosí espresse in forma covariante per mezzo di un solo tensore emisimmetrico.

Quanto alle forze agenti sulla carica Q in moto, secondo Lorentz si ha

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = Q\left(\mathbf{E} + \frac{\mathbf{v} \wedge \mathbf{B}}{c}\right) \tag{4.43}$$

che, moltiplicata per $1/\sqrt{1-\beta^2}$, assume la forma covariante

$$\frac{dp^\alpha}{du} = \frac{Q}{c^2} F^{\alpha\mu} u_\mu \tag{4.44}$$

dove u/c è il tempo proprio. La componente temporale della (4.44) si riduce a $dE/dt = Q(\mathbf{v} \cdot \mathbf{E})$, ed esprime la variazione di energia di Q per unità di tempo come lavoro della forza agente su Q.

Si ha dunque una serie completa di equazioni relativisticamente invarianti che risolvono il problema del moto di cariche e CEM accoppiati. Tuttavia una discussione esauriente dovrebbe necessariamente includere anche una derivazione lagrangiana di queste equazioni [il lettore interessato la troverà al § 31 di Pauli 1921].

5. *La teoria della relatività generale: introduzione geometrica.*

5.1. Il principio di equivalenza debole.

La teoria del campo gravitazionale poggia sul principio di equivalenza tra gravitazione e inerzia. Molti degli effetti fisici della teoria della relatività generale possono già essere dimostrati o intuiti partendo da questo principio. Lo sviluppo completo e le ultime conseguenze del principio possono dedursi solamente attraverso il calcolo tensoriale.

Fu già osservato da Galileo che i corpi cadono con un'accelerazione che non dipende dalla loro massa. L'osservazione di Galileo fu ampliata e controllata con crescente grado di precisione da Huygens, Newton, Bessel, fino al barone Eötvös che pose stringenti limiti sperimentali alle deviazioni dal principio [1889]. Nella teoria non relativistica della gravitazione newtoniana è possibile formalizzare facilmente il cosiddetto principio di equivalenza debole. Già è noto che tra l'accelerazione a di un corpo e la forza F che agisce su di esso vale la relazione $F = m_i a$, dove la massa inerziale m_i è un parametro caratteristico del corpo. In particolare, se la forza è di origine gravitazionale, è noto che tra il corpo e, ad esempio, la Terra, la forza è $F = G m_g M_g / r^2$, dove m_g e M_g sono altri parametri, detti masse gravitazionali del corpo e della Terra rispettivamente. Ne segue che l'accelerazione di gravità g deve valere

$$g = G \frac{m_g}{m_i} \frac{M_g}{r^2}.$$

L'esperienza di Galileo e altri dice che, con grande accuratezza,

$$m_g = m_i, \tag{5.1}$$

cosicché g dipende solamente dai parametri relativi alla Terra e non dal corpo. La (5.1) è nota come principio di equivalenza debole e verrà qui assunta come punto di partenza delle speculazioni che seguono.

Einstein mostrò in una famosa esperienza concettuale che, valendo questo principio, è impossibile in generale mettere in evidenza il campo gravitazionale esterno stando entro un ascensore in caduta libera. Infatti tutti i corpi racchiusi entro la cabina cadrebbero esattamente con la stessa accelerazione dell'ascensore, e non sarebbe possibile scoprire alcuna accelerazione relativa. Inversamente, se si accelera una cabina nello spazio esterno, lontano da ogni massa, si ottiene entro la cabina lo stesso effetto di un campo gravitazionale esterno costante. Questa possibilità di trasformare l'accelerazione di gravità in inerzia, e viceversa, costituisce il punto di partenza del principio di equivalenza nella sua versione forte. Esso all'inizio altro non è che un manifesto scientifico, la cui portata reale diverrà chiara solamente entro una precisa cornice matematica.

Prima di formulare e discutere il principio di equivalenza, conviene quindi anticipare alcune notazioni e una terminologia che troveranno impiego e completa giustificazione in quanto segue.

5.2. Carte locali, varietà differenziabili.

Concetto basilare è quello di spazio topologico di Hausdorff **W**.

DEFINIZIONE 1. *Uno spazio di Hausdorff* **W** *è un insieme di elementi (punti) tale che esiste una collezione* Σ *di sottoinsiemi di* **W** *(detti aperti di* **W***) che soddisfa:*

1) *Una unione di aperti, anche infinita, è ancora un aperto.*
2) *Una intersezione finita* $\bigcap_{i=1}^{n<\infty} \mathcal{U}_i \in \Sigma$.
3) *Dati due punti qualsiasi* $x_1, x_2 \in \mathbf{W}$, *esistono in* Σ *aperti disgiunti* $\mathcal{U}_1, \mathcal{U}_2$, *tali che* $x_1 \in \mathcal{U}_1$, $x_2 \notin \mathcal{U}_1$, $x_2 \in \mathcal{U}_2$, $x_1 \notin \mathcal{U}_2$.

Gli assiomi 1), 2) caratterizzano uno spazio topologico qualunque; l'assioma 3) è l'assioma di separazione di Hausdorff. La collezione Σ si dice topologia di **W**. Nello spazio $\mathbf{R}^n$ la topologia Σ consiste di tutti gli aperti generati tramite le condizioni 1), 2), a partire dagli aperti cubici del tipo: $a^\alpha < x^\alpha < b^\alpha$, $\alpha = 1, \ldots, n$. Ad esempio, l'interno della sfera $\sum_\alpha |x^\alpha|^2 < 1$ è un aperto.

DEFINIZIONE 2. *Una carta aperta locale* $(\mathcal{U}^\alpha, \mathcal{C})$ *è formata da un aperto* $\mathcal{U} \in \Sigma$ *e da un omeomorfismo* $\mathcal{C} : \mathcal{U} \to \mathcal{O}$, *dove* $\mathcal{O}$ *è un aperto di* $\mathbf{R}^n$.

DEFINIZIONE 3. *Una struttura differenziale su* **W** *di dimensione* m *è una collezione* $\mathcal{A}$ *di carte, detta atlante,* $(\mathcal{U}_a, \mathcal{C}_a)$, $a \in \mathcal{A}$, *dove* $C_a(\mathcal{U}_a)$ *è un aperto in* $\mathbf{R}^m$ *tale che:*

1) $\bigcup_a \mathcal{U}_a = \mathbf{W}$.
2) *Se* $a, b \in \mathcal{A}$, $\mathcal{U}_a \cap \mathcal{U}_b \neq \varnothing$ *(*$\varnothing$ *insieme vuoto), l'applicazione* $\mathcal{C}_b \mathcal{C}_a^{-1}$ *è un'applicazione differenziabile di* $\mathcal{C}_a(\mathcal{U}_a \cap \mathcal{U}_b)$ *su* $\mathcal{C}_b(\mathcal{U}_a \cap \mathcal{U}_b)$.

In linguaggio intuitivo, la $\mathcal{C}_a$ assegna coordinate locali $x_a^1, \ldots, x_a^m$ in $\mathcal{U}_a$, e $\mathcal{C}_b$ assegna $x_b^1, \ldots, x_b^m$ in $\mathcal{U}_b$. In $\mathcal{U}_a \cap \mathcal{U}_b$ si può usare sia $x_a^1, \ldots, x_a^m$, sia $x_b^1, \ldots, x_b^m$ per lo stesso punto x. La 2) richiede che la corrispondenza tra x_a^α, x_b^β sia differenziabile ed invertibile, cioè lo jacobiano $\det \mathcal{J}_{ab} \equiv |\mathcal{J}_{ab}|$, $\mathcal{J}_{ab} \equiv \partial(x_a^1, \ldots, x_a^m)/\partial(x_b^1, \ldots, x_b^m)$, deve essere finito, non nullo. Si assumerà inoltre che tutte le applicazioni differenziabili di cui al 2) siano lisce, cioè infinitamente differenziabili o, come si dice, di classe C^∞. Se $m = 0$, s'intenderà per $\mathbf{R}^0$ un singolo punto.

DEFINIZIONE 4. *Una varietà differenziabile* **W** *di dimensione* m *è uno spazio di Hausdorff dotato di struttura differenziabile* $\mathcal{A}$ *di dimensione* m. *Quindi una coppia* $(\mathcal{U}_a, \mathcal{C}_a)$, *dove* $a \in \mathcal{A}$, *definisce un sistema locale di coordinate su* **W**: *al punto* $x \in \mathbf{W}$, *essa assegna coordinate* $x_a^1, \ldots, x_a^m$.

DEFINIZIONE 5. *Una funzione reale* $f : \mathbf{W} \to \mathbf{R}$ *dicesi differenziabile in* $x \in \mathbf{W}$ *se* $f\,\mathcal{C}_a^{-1}$ *è differenziabile in* $\mathcal{C}_a(\mathcal{U}_a)$ *per* $x \in \mathcal{U}_a$. f *dicesi differenziabile se lo è in tutti i punti* $x \in \mathbf{W}$.

Ciò significa, in sostanza, che se f è differenziabile in x, essa è differenziabile nelle x_a^μ, x_b^ν, $\forall a, b$ tali che $x \in \mathcal{U}_a \cap \mathcal{U}_b$. In quanto segue s'intenderà sempre differenziabilità di classe C^∞. L'insieme di tutte le funzioni differenziabili su **W** sarà indicato $C^\infty(\mathbf{W})$.

DEFINIZIONE 6. *Sia* $\mathbf{W}^n$ *differenziabile e* $\mathcal{A}$ *un atlante di* $\mathbf{W}^n$. *Per orientazione della carta* $(\mathcal{U}_a, \Phi_a) \in \mathcal{A}$ *s'intenderà un ordine specifico delle variabili* $x_a^1, \ldots, x_a^n$, *assegnato a meno di permutazioni pari di queste variabili. Quindi, scambiando due coordinate, si ottiene l'orientazione opposta. Siano* $(\mathcal{U}_a, \Phi_a)$, $(\mathcal{U}_b, \Phi_b)$ *due carte in* $\mathcal{A}$ *e* $\mathcal{U}_a \cap \mathcal{U}_b \neq \varnothing$. *In* $\mathcal{U}_a \cap \mathcal{U}_b$ *esiste una corrispondenza tra* $x_a^1, \ldots, x_a^n$ *e* $x_b^1, \ldots, x_b^n$ *che è differenziabile ed invertibile, e cioè* $|\mathcal{J}_{ab}| \neq 0$, $|\mathcal{J}_{ab}|$ *finito. Se* $|\mathcal{J}_{ab}| > 0$, *si dirà che* $\mathcal{U}_a$, $\mathcal{U}_b$ *hanno la stessa orientazione, se* $|\mathcal{J}_{ab}| < 0$, *si dirà che hanno opposta orientazione. Se è possibile ordinare le variabili in ciascuna carta in modo da avere* $|\mathcal{J}_{ab}| > 0$ $\forall a, b$, *si dirà che* $\mathbf{W}^n$ *è orientabile.*

DEFINIZIONE 7. *Una varietà* $\mathbf{V}^m$ *è una sottovarietà di* $\mathbf{W}^n$ *(dove* $m \leq n$*) se esiste un'applicazione* $\xi : \mathbf{V}^m \to \mathbf{W}^n$ *regolare, cioè se esiste una carta* $(\mathcal{U}_a, \Phi_a)$ *di* $\mathbf{W}^n$ *e una* $(\mathcal{U}'_b, \Psi_b)$ *di* $\mathbf{V}^m$ *tali che la matrice* $\mathcal{J}$, *rettangolare* $m \times n$, *di elementi* $\partial y_b^i / \partial x_a^k$ *ha rango massimo* m, *ovvero esiste almeno una sottomatrice* $m \times m$, *tra le* $\binom{n}{m}$ *estraibili da* $\mathcal{J}$, *che ha determinante non nullo. In questo caso si dice che* $\mathbf{V}^m$ *è immersa in* $\mathbf{W}^n$.

Ad esempio, tutti gli iperpiani $\mathbf{R}^{n-1}$ in $\mathbf{R}^n$ sono sottovarietà di $\mathbf{R}^n$. La sfera $S^{n-1} \subset \mathbf{R}^n$ data da $\sum_{i=1}^n (x^i)^2 = 1$, è anche una sottovarietà di $\mathbf{R}^n$. Una linea $x^i = f^i(t)$ è una $\mathbf{V}^1 \subset \mathbf{W}^n$ purché per ogni t esista i tale che $dx^i/dt \neq 0$; in questo senso, la $x = t^2$, $y = t^3$, nel piano non è una sottovarietà, per via della singolarità cuspidale in $t = 0$.

5.3. Il principio di equivalenza.

Il principio di equivalenza, nella sua forma forte, stabilisce che in ogni evento dello spazio-tempo è possibile scegliere una carta locale di coordinate, detta localmente inerziale, tale che, in un intorno sufficientemente piccolo dell'evento, le leggi di natura si scrivano nello stesso modo in cui appaiono nella relatività ristretta [Einstein 1916].

Cosa s'intenda per «stesso modo» e la grandezza di questo intorno sarà precisato dal calcolo tensoriale. Dovrebbe essere evidente che per intorno dell'evento s'intendono gli eventi interni all'ascensore di Einstein, in cui il campo gravitazionale sia stato annullato dalla caduta libera. La portata del principio è molto piú ampia; l'esperimento concettuale di Einstein si riferisce dapprima solamente a gravi in caduta libera non relativistica. Il principio richiede invece l'annullarsi esatto di tutti gli effetti di gravità per qualunque sistema fisico, ivi incluso il campo elettromagnetico.

E dovrebbe essere chiaro anche che questa compensazione del campo gravitazionale può essere esatta solamente in un evento (o meglio lungo una linea di eventi nello spazio-tempo), ma non in una regione finita a quattro dimensioni nello spazio-tempo. Infatti se il campo gravitazionale non è omogeneo, non è pos-

sibile compensare accelerazioni diverse in diversi punti della cabina con una sola accelerazione globale. Rimarranno forze residue, della stessa natura di quelle che dànno origine alle maree sulla Terra. Il principio di equivalenza è dunque un concetto limite, valido a rigore solamente in un intorno infinitesimo di un dato evento.

Siano P_a e P_b due eventi vicini nello spazio-tempo. Per quanto detto, ognuno di questi eventi avrà una carta $(\mathcal{U}_i, \Phi_i)$ dove $i=a,b$, e tale che in P_i si ha

$$(5.2) \qquad -du^2 = \eta_{\mu\nu}\, dx_i^\mu\, dx_i^\nu$$

(seguendo la consuetudine, $\eta_{\mu\nu}$ indica il tensore metrico della relatività ristretta, cioè per uno spazio senza curvatura, mentre $g_{\mu\nu}$ sarà riservato al caso con curvatura). Ciascuna delle carte $\mathcal{U}_i$ è dunque inerziale in P_i. L'esistenza di un campo gravitazionale fa sí che il sistema localmente inerziale in P_a e quello inerziale in P_b siano dotati di moto relativo *non* uniforme. In $\mathcal{U}_a \cap \mathcal{U}_b$ la relazione tra le coordinate x_a^μ e x_b^μ non può essere lineare, le derivate $\partial x_a^\mu / \partial x_b^\nu$ non sono dunque costanti. Questa non-linearità riflette essenzialmente la mancata compensazione del campo gravitazionale in P_a quando il sistema sia esattamente inerziale in P_b, secondo quanto accennato poco prima.

La metrica in P_a, espressa mediante le coordinate x_b^ν, appare nella forma

$$(5.3) \qquad -du^2 = \eta_{\mu\nu} \frac{\partial x_a^\mu}{\partial x_b^\lambda} \frac{\partial x_a^\nu}{\partial x_b^\sigma}\, dx_b^\lambda\, dx_b^\sigma \equiv g_{\mu\nu,b}\, dx_b^\mu\, dx_b^\nu,$$

che non ha coefficienti costanti; si ha dunque: $g_{\mu\nu,b} \neq \eta_{\mu\nu}$. Il tentativo di estendere la metrica $g_{\mu\nu,b}$ lontano dal punto P_b conduce a una dipendenza di $g_{\mu\nu,b}$ dalle coordinate. Se ne deduce che, in presenza del campo gravitazionale, non è possibile rappresentare ovunque la metrica usando la formula di Minkowsky.

Questo conduce in modo naturale allo schema seguente:

A) $\mathfrak{M}$ è una varietà differenziabile dotata di atlante.

B) In ogni carta è data una metrica $g_{\mu\nu,a}$ tale che in $\mathcal{U}_a \cap \mathcal{U}_b$

$$(5.4) \qquad g_{\mu\nu,b} = g_{\lambda\sigma,a} \frac{\partial x_a^\lambda}{\partial x_b^\mu} \frac{\partial x_a^\sigma}{\partial x_b^\nu}.$$

C) In una carta generica si ha $g_{\mu\nu,a} \neq \eta_{\mu\nu}$; pertanto si richiede che sia possibile scegliere una carta nell'intorno di un evento dato, in cui valga l'inerzialità (5.2).

Secondo quanto sarà sviluppato in seguito, la condizione C) è derivabile dalle altre. A) e B) contengono dunque già incorporato il principio di equivalenza.

5.4. Varietà riemanniane: curvatura, trasporto parallelo.

Una varietà dotata della struttura addizionale di cui in A), B), C), sarà detta varietà riemanniana, da Bernhard Riemann che per primo le definí e ne intuí le

importantissime proprietà. In una varietà di Riemann, dunque, esiste una formula per il tempo proprio data in una carta generica $\mathcal{U}_a$ da

$$(5.5) \qquad -du^2 = g_{\mu\nu,a}\, dx_a^\mu\, dx_a^\nu.$$

Il primo contributo allo studio di tali varietà si ebbe con la pubblicazione del trattato *Disquisitiones generales circa superficies curvas* di Gauss [1828], dove furono poste le basi della geometria differenziale. Tuttavia neppure Gauss poté certo immaginare che il suo lavoro sulle varietà bidimensionali, una volta generalizzato a quattro, potesse servire a spiegare la relatività.

Il passo seguente fu compiuto appunto da Riemann con il suo famoso lavoro *Sulle ipotesi che stanno alla base della Geometria* [1854], dove venne generalizzato il lavoro di Gauss a piú dimensioni e fondata la geometria riemanniana. Riemann fece tuttavia qualcosa di piú; egli considerò la varietà come un'entità a sé stante, contrariamente a quanto aveva fatto Gauss, il cui punto di partenza erano le superfici immerse in $\mathbf{R}^3$. Riemann liberò le coordinate dal loro significato metrico, riducendole a una funzione etichettatrice dei punti della varietà: fintantoché non viene assegnato il tensore metrico $g_{\mu\nu,a}$ in una carta, le coordinate x_a non dànno alcuna informazione sulle distanze dei punti che rappresentano. Che questo punto di vista costituisse un sostanziale passo avanti, di notevole portata, è ammesso dallo stesso Einstein.

Il passaggio alle coordinate generiche fece sorgere il primo problema fondamentale della geometria riemanniana: date due metriche $g_{\mu\nu,a}$ e $g_{\lambda\sigma,b}$, è possibile riconoscere se esse sono legate da una trasformazione (5.4) e, quindi, se rappresentano la stessa varietà? Questo problema fu risolto da Christoffel [1869] e Lipschitz [1869]. Basandosi sui risultati di questi matematici, Ricci nel 1887-1896 sviluppò l'analisi tensoriale senza la quale la relatività non avrebbe mai conosciuto lo sviluppo attuale. Levi-Civita, allievo di Ricci, approfondí ulteriormente i concetti, introducendo il trasporto parallelo [cfr. Ricci e Levi-Civita 1901; Levi-Civita 1917]. Se lo sviluppo della fisica deve molto a questi studiosi, l'avvento della relatività generale nel 1916 spronò l'attività dei matematici.

Un'ulteriore generalizzazione dei concetti riemanniani avvenne con il lavoro di Cartan *Sur les variétés à connexion affine et la théorie de la Relativité Généralisée* [1923-24]. Nessuno dei fisici, e forse neppure Einstein, comprese la portata del lavoro di Cartan. Questo era in parte dovuto all'uso delle forme differenziali in luogo del calcolo tensoriale, piú familiare ai fisici. Oltre a contributi fondamentali alla geometria differenziale, riconosciuti tali solo dopo molti anni, Cartan introdusse i concetti di torsione e di connessione, e pose le basi per le moderne teorie unificate.

Tornando al concetto di metrica, ci si propone di discuterne la portata con alcuni semplici esempi, nei quali si supporrà che la metrica sia definita positiva. È evidente, per quanto detto, che le varietà a una dimensione sono tutte riducibili a $\mathbf{R}$ oppure a S^1, in quanto esiste una sola coordinata; la metrica si riduce quindi a $du^2 = g_a\, dx_a^2$; ponendo $x_b = \int \sqrt{g_a}\, dx_a$, si giunge a $du^2 = (dx_b)^2$, dove $x_b \in \mathbf{R}$, oppure $x_b \in S^1$ secondo il caso; non esiste quindi geometria intrinseca di rilievo in una dimensione.

Di fatto la curvatura di una linea è un concetto relativo all'immersione di questa in uno spazio piú grande. Il raggio di curvatura è definito nel modo seguente. Sia u un parametro monotòno lungo una curva Γ e siano $x(u)$, $y(u)$ le funzioni parametriche di Γ nel piano x, y. Si ha $du^2 = dx^2 + dy^2$; quindi $(dx/du)^2 + (dy/du)^2 = 1$, per cui si può porre $dx/du = -\sin\varphi$, $dy/du = \cos\varphi$; la curvatura ρ è quindi definita da $\rho = d\varphi/du$. Infatti per una circonferenza, se $x = R\cos(u/R)$, $y = R\sin(u/R)$, si ha $\varphi = u/R$, $\rho = 1/R$. È possibile definire intuitivamente ρ alzando due normali a Γ in due punti $P(u_1)$, $P(u_2)$ di Γ. Sia $\Delta\varphi$ l'angolo fra le due normali; $\rho = \lim_{u_1 \to u_2} \Delta\varphi/(u_1 - u_2)$. La normale a Γ in $P(u)$ definisce un punto su di un cerchio S^1 di coordinata φ; esiste dunque un'applicazione $\Gamma \to S^1$, e la curvatura è il rapporto tra archi corrispondenti in S^1 e in Γ quando si faccia il limite indicato.

In modo del tutto analogo Gauss considerò su di una superficie generica Σ l'insieme dei versori normali a Σ in ciascun punto. Questi, trasportati in un unico punto di applicazione O, descrivono punti di una sfera S^2 di raggio 1 e individuano un'applicazione $\gamma : \Sigma \to S^2$ che estende quella data per le curve. Sia σ un piccolo insieme su Σ e $P \in \sigma$; sia poi $\gamma(\sigma)$ l'immagine di σ su S^2. La quantità

$$K_P = \lim_{\sigma \to P} \frac{\text{area}(\gamma(\sigma))}{\text{area}(\sigma)} \tag{5.6}$$

fu presa da Gauss come definizione di curvatura di Σ in P.

La definizione (5.6) non è intrinseca, dato che per applicarla bisogna per forza immergere Σ in $\mathbf{R}^3$. Tuttavia, con un vero *tour de force*, Gauss riuscí a dimostrare che K_P poteva essere calcolato dalla metrica di Σ dedotta nel modo seguente. Sia $x(t,v)$, $y(t,v)$, $z(t,v)$ una parametrizzazione arbitraria di Σ; la distanza tra punti vicini sarà data da $dx^2 + dy^2 + dz^2 = E\,dt^2 + 2F\,dv\,dt + G\,dv^2 = g_{\mu\nu}dx^\mu dx^\nu$, ove si ponga $t = x^1$, $v = x^2$, e

$$E \equiv g_{11} = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2$$

$$F \equiv g_{12} = g_{21} = \frac{dx}{dt}\frac{dx}{dv} + \frac{dy}{dt}\frac{dy}{dv} + \frac{dz}{dt}\frac{dz}{dv}$$

$$G \equiv g_{22} = \left(\frac{dx}{dv}\right)^2 + \left(\frac{dy}{dv}\right)^2 + \left(\frac{dz}{dv}\right)^2.$$

I simboli E, F, G sono quelli usati originalmente da Gauss, che riuscí ad esprimere K in funzione di E, F, G, e delle loro derivate prime e seconde.

Due superfici che abbiano la stessa K, possono flettersi fino a sovrapporsi esattamente. Ad esempio, sia il piano sia il cilindro hanno $K = 0$, e si può sempre immaginare il cilindro come un piano avvolto infinite volte. Non si può invece svolgere la sfera e adagiarne la superficie su di un piano; ciò è impossibile anche per una calotta sferica. Altrimenti detto, l'unico parametro che regola realmente la geometria di Σ è la curvatura, che non dipende dalla scelta delle coordinate x^1, x^2.

Per una sfera di raggio R, $K = 1/R^2$, risultato che segue direttamente da (5.6) e che suggerisce un'altra definizione di K, intrinseca e vicinissima a quella moderna. Sulla sfera vale un noto teorema di trigonometria, secondo il quale l'area A di un triangolo sferico avente angoli interni α, β, γ, è data da $A/R^2 = \alpha + \beta + \gamma - \pi$, da cui si ottiene

$$K = \frac{1}{R^2} = \frac{\alpha + \beta + \gamma - \pi}{A}; \tag{5.7}$$

si noti che la (5.7) è scritta totalmente in termini di geometria intrinseca, poiché le distanze, gli angoli e le aree non richiedono il concetto di immersione. Basandosi sulla (5.7), è possibile definire K nel modo seguente.

Si prendano tre punti P_1, P_2, P_3 molto vicini in Σ e si costruiscano gli archi di curva piú brevi (geodetiche) congiungenti detti punti a coppie in modo da formare un triangolo analogo a quello sferico; indi si definisca

$$K = \lim_{A \to 0} \frac{\alpha_1 + \alpha_2 + \alpha_3 - \pi}{A} \tag{5.8}$$

dove α_i è l'angolo interno tra le geodetiche in P_i, $i = 1, 2, 3$, e A è l'area del triangolo iniziale. Come esempi di geodetiche si possono ricordare le linee rette nel piano e i circoli massimi sulle sfere. Esistono superfici, le pseudosfere, per le quali $K = \text{cost} < 0$ e quindi $\alpha_1 + \alpha_2 + \alpha_3 < \pi$. In generale, K non è costante su Σ; essa risulta positiva in $P \in \Sigma$ quando Σ sta tutta da una parte rispetto al piano tangente a Σ in P, negativa quando questo attraversa Σ (P punto a sella).

Per mezzo del calcolo tensoriale, che verrà sviluppato nel prossimo paragrafo, la (5.8) sarà opportunamente generalizzata. Si considererà infine una variante della (5.8) che introdurrà intuitivamente il concetto di trasporto parallelo. Sia P_1, P_2, P_3 un triangolo geodetico su Σ e gli archi siano orientati come (P_1P_2), (P_2P_3), (P_3P_1); si definirà trasporto parallelo di un vettore $\mathbf{v}$ lungo un arco di geodetica una famiglia di vettori $\mathbf{v}(s)$, dove s è un parametro lungo la geodetica, $|\mathbf{v}(s)| = |\mathbf{v}|$, tali che $\mathbf{v}(s)$ formi, con la tangente alla geodetica, un angolo costante al variare di s. Senza perdere in generalità, si parte da P_1, e si sceglie tale angolo uguale a zero; quindi $\mathbf{v}(s)$ riproduce progressivamente tutti i vettori tangenti a (P_1P_2). Il vettore, cosí trasportato, quando giunge in P_2 forma un angolo $\pi - \alpha_2$ con il versore tangente in P_2 a (P_2P_3) e, trasportato lungo (P_2P_3), giunge in P_3 con un angolo $2\pi - \alpha_2 - \alpha_3$ rispetto al versore tangente in P_3 a (P_3P_1). Finalmente, ritorna in P_1 formando con il versore tangente in P_1 a (P_1P_2) (cioè il $\mathbf{v}$ di partenza) un angolo pari a $3\pi - \alpha_1 - \alpha_2 - \alpha_3 = \pi - \alpha_1 - \alpha_2 - \alpha_3 = -KA \neq 0$. Pertanto da questo semplice esempio si vede che il trasporto parallelo di un vettore lungo una linea chiusa dà lo stesso vettore ruotato di un ammontare proporzionale alla curvatura integrata sull'area racchiusa dal cammino. Questo concetto sarà formalizzato nel trasporto parallelo di Levi-Civita.

Lo schema A), B), C) appare dunque come una generalizzazione degli esempi discussi finora al caso di quattro dimensioni. Occorre tuttavia precisare che il lavoro originale di Riemann, e gran parte di quelli di geometria differenziale che ne sono seguiti, si basano su una metrica definita positiva; nella relatività invece

la metrica possiede tre autovalori positivi di tipo spaziale e uno negativo di tipo temporale.

Dagli esempi considerati è chiaro che non è possibile scegliere su di una generica varietà riemanniana una carta $(\mathfrak{U}_a, \Phi_a)$ in cui la (5.2) sia verificata in tutto l'insieme $\mathfrak{U}_a$, proprio come non è possibile scegliere coordinate x_1, x_2, sulla sfera in modo da avere $du^2 = (dx_1)^2 + (dx_2)^2$. Altrimenti, la geometria intrinseca della sfera sarebbe esattamente identica a quella del piano, e sarebbe possibile disegnare atlanti del globo fedeli nelle distanze, aree ed angoli. Dunque le generiche $g_{\mu\nu}$ definiscono una struttura metrica che non è equivalente a quella dello spazio-tempo della relatività ristretta. Questa struttura contiene tutte le informazioni necessarie ad identificare il campo gravitazionale; in questo senso, sarà identificata con il campo gravitazionale.

6. *Il calcolo tensoriale su una varietà liscia.*

6.1. Proprietà generali.

Sia $\mathfrak{M}^n$ una varietà liscia supposta, per maggiore generalità, a n dimensioni; il caso $n=4$ sarà pertinente alle applicazioni alla relatività generale. $\mathfrak{M}^n$ possiede una struttura metrica; nella carta $(\mathfrak{U}_a, \mathcal{C}_a)$ si ha:

$$ds^2 = g_{\mu\nu,a}\, dx_a^\mu dx_a^\nu. \tag{6.1}$$

Per tensore misto del tipo (q, p) s'intenderà una collezione di elementi

$$T^{\mu_1,\mu_2,\ldots,\mu_q}_{\nu_1,\nu_2,\ldots,\nu_p;\,a}(x) \tag{6.2}$$

assegnati in ciascuna carta $(\mathfrak{U}_a, \mathcal{C}_a)$, dove $\mu_i = 1, \ldots, n$; $\nu_i = 1, \ldots, n$, e soggetti alle condizioni di consistenza in $\mathfrak{U}_a \cap \mathfrak{U}_b \neq \varnothing$:

$$T^{\mu_1,\ldots,\mu_q}_{\nu_1,\ldots,\nu_p;\,a}(x) = \frac{\partial x_a^{\mu_1}}{\partial x_b^{\lambda_1}} \cdots \frac{\partial x_a^{\mu_q}}{\partial x_b^{\lambda_q}} \frac{\partial x_b^{\sigma_1}}{\partial x_a^{\nu_1}} \cdots \frac{\partial x_b^{\sigma_p}}{\partial x_a^{\nu_p}}\, T^{\lambda_1,\ldots,\lambda_q}_{\sigma_1,\ldots,\sigma_p;\,b}(x). \tag{6.3}$$

Osservazione 1: μ_i sarà detto indice controvariante, ν_i covariante. T^μ si dirà vettore controvariante, T_ν vettore covariante e se $q = p = 0$, T si dirà scalare. Si noti che, poiché $dx_a^\mu = (\partial x_a^\mu / \partial x_b^\lambda)\, dx_b^\lambda$, dx_a^μ risulta essere un vettore controvariante. Questa nomenclatura ricalca quella già usata in relatività ristretta e ad essa si riduce se ci si limita a carte inerziali in $\mathfrak{M}$. La differenza sostanziale è che, nella (6.3), le $\partial x_a^\mu / \partial x_b^\nu$ non sono in generale costanti, né $g_{\mu\nu,a}$ è costante, né occorre che $g_{\mu\nu,a}$ sia lo stesso per tutte le carte. Nella relatività ristretta ci si limitava a carte inerziali nelle quali $g_{\mu\nu,a} = \eta_{\mu\nu}$; ciò poneva forti restrizioni sulle carte, cioè $\partial x_a^\mu / \partial x_b^\nu$ erano matrici di Lorentz. Come detto, queste restrizioni non sussistono per la (6.3). Ciò nondimeno valgono quasi tutte le regole del calcolo tensoriale già enunciate nel § 3.2, ad eccezione della regola 5, che necessita modificazioni di altissimo valore concettuale formanti la parte piú qualificante dell'opera di Ricci e Levi-Civita.

Osservazione 2: Le $\delta^\mu_{\nu,a} = \delta^\mu_\nu$ (simbolo di Kronecker), con $\delta^\mu_\nu = 0$ se $\mu \neq \nu$, $\delta^\mu_\nu = 1$ se $\nu = \mu$, dati in ogni carta, formano un tensore di tipo $(1, 1)$. Infatti, dalla (6.3),

$$\delta^\lambda_\sigma \frac{\partial x_a^\mu}{\partial x_b^\lambda} \frac{\partial x_b^\sigma}{\partial x_a^\nu} = \frac{\partial x_a^\mu}{\partial x_b^\lambda} \frac{\partial x_b^\lambda}{\partial x_a^\nu} = \delta^\mu_\nu,$$

in virtú delle regole di derivazione. Si noti inoltre che applicando $\partial / \partial x_a^\sigma$ a quest'ultima identità si ottiene

$$\frac{\partial^2 x_a^\mu}{\partial x_b^\tau \partial x_b^\lambda} \frac{\partial x_b^\tau}{\partial x_a^\sigma} \frac{\partial x_b^\lambda}{\partial x_a^\nu} = -\frac{\partial x_a^\mu}{\partial x_b^\lambda} \frac{\partial^2 x_b^\lambda}{\partial x_a^\sigma \partial x_a^\nu}.$$

Osservazione 3: Converrà introdurre le quantità $g_a^{\mu\nu}$, che formano la matrice inversa alla $g_{\mu\nu,a}$: $g_a^{\mu\nu} g_{\nu\lambda,a} = \delta^\mu_\lambda$. Come nel § 3, verrà indicata con g_a la matrice formata dalle $g_{\mu\nu,a}$ e con g_a^{-1} l'inversa. Inoltre s'introdurrà la matrice jacobiana $\mathcal{J}_{ab}$ di elementi $\mathcal{J}_{ab\lambda}^{\ \ \mu} = \partial x_a^\mu / \partial x_b^\lambda$ e la $\mathcal{J}_{ba}$ di elementi $\partial x_b^\mu / \partial x_a^\lambda$, dove $\mathcal{J}_{ab} = \mathcal{J}_{ba}^{-1}$. La (6.3), applicata a g_a, dà

$$g_a = (\mathcal{J}_{ba})^T g_b \mathcal{J}_{ba} \tag{6.4}$$

e, invertendo, $g_a^{-1} = \mathcal{J}_{ab} g_b^{-1} (\mathcal{J}_{ab})^T$, da cui

$$g_a^{\mu\nu} = \frac{\partial x_a^\mu}{\partial x_b^\sigma} \frac{\partial x_a^\nu}{\partial x_b^\tau} g_b^{\sigma\tau}. \tag{6.5}$$

Osservazione 4: Operando in una carta data a, si può utilizzare il tensore metrico $g^{\mu\nu}$ o $g_{\alpha\beta}$ per alzare o abbassare un indice secondo le regole

$$T^{\mu(\sigma)}_{(\lambda)} = g^{\mu\nu} T^{(\sigma)}_{\nu(\lambda)} \qquad T^{(\sigma)}_{\nu(\lambda)} = g_{\nu\mu} T^{\mu(\sigma)}_{(\lambda)}$$

nelle quali è stato tralasciato l'indice a; (σ) e (λ) sono insiemi di indici tensoriali che rimangono immutati.

Si possono formare nuovi tensori mediante le regole già enunciate nel § 3, che verranno ripetute qui, omettendo l'indicazione della carta.

1) Se $T^{\mu_1,\ldots,\mu_q}_{\nu_1,\ldots,\nu_p}$, $S^{\mu_1,\ldots,\mu_q}_{\nu_1,\ldots,\nu_p}$ sono tensori dello stesso tipo, le quantità $R^{\mu_1\ldots\mu_q}_{\nu_1\ldots\nu_p} = \alpha T^{\mu_1\ldots\mu_q}_{\nu_1\ldots\nu_p} + \beta S^{\mu_1\ldots\mu_q}_{\nu_1\ldots\nu_p}$, α e β scalari, formano un tensore dello stesso tipo.
2) Se $T^{\mu_1\ldots\mu_q}_{\nu_1\ldots\nu_p}$ è di tipo (q, p), $R^{\mu_1\ldots\mu_{q-1}}_{\nu_1\ldots\nu_{p-1}} = T^{\mu_1\ldots\mu_{q-1}\alpha}_{\nu_1\ldots\nu_{p-1}\alpha}$ è un tensore di tipo $(q-1, p-1)$.
3) Se $T^{\mu_1\ldots\mu_q}_{\nu_1\ldots\nu_p}$, $S^{\mu_1\ldots\mu_s}_{\nu_1\ldots\nu_t}$ sono di tipo (q, p), (s, t), allora $R^{\mu_1\ldots\mu_{q+s}}_{\nu_1\ldots\nu_{p+t}} = T^{\mu_1\ldots\mu_q}_{\nu_1\ldots\nu_p} S^{\mu_{q+1}\ldots\mu_{q+s}}_{\nu_{p+1}\ldots\nu_{p+t}}$ è un tensore di tipo $(s+q, t+p)$.
4) Se $T^{\mu_1\ldots\mu_q}_{\nu_1\ldots\nu_p}$ è un tensore di tipo (q, p), formano un tensore (q, p) le quantità $R^{\mu_1\ldots\mu_q}_{\nu_1\ldots\nu_p} = T^{\mu_{\sigma(1)}\ldots\mu_{\sigma(q)}}_{\nu_{\tau(1)}\ldots\nu_{\tau(p)}}$, dove $\sigma(i)$, $\tau(i)$ sono permutazioni arbitrarie fisse dei numeri $1, \ldots, q$; $1, \ldots, p$, rispettivamente.

Non vale invece la regola 5 di derivazione, nella sua formulazione originale, per la seguente ragione. Sia T^μ un vettore controvariante; applicando $\partial / \partial x_a^\nu$ alla $T_a^\mu = (\partial x_a^\mu / \partial x_b^\lambda) T_b^\lambda$, si ha

$$\frac{\partial T_a^\mu}{\partial x_a^\nu} = \frac{\partial x_a^\mu}{\partial x_b^\lambda} \frac{\partial x_b^\sigma}{\partial x_a^\nu} \frac{\partial T_b^\lambda}{\partial x_b^\sigma} + \frac{\partial^2 x_a^\mu}{\partial x_b^\lambda \partial x_b^\sigma} \frac{\partial x_b^\sigma}{\partial x_a^\nu} T_b^\lambda;$$

il primo contributo a destra corrisponde alla (6.3) applicata ad un tensore di tipo (1, 1), mentre il secondo contributo esce dalla definizione di tensore, per cui $\partial T^{\mu}/\partial x^{\nu}$ non è un tensore.

Questo fatto ha costituito un ostacolo nei primi anni di ricerca sulle varietà di Riemann. Per ovviarlo fu creato il calcolo differenziale assoluto [cfr. Ricci e Levi-Civita 1901; Levi-Civita 1917].

6.2. Il trasporto parallelo e i coefficienti di connessione.

Per trovare le regole di derivazione covariante in varietà con curvatura, occorre utilizzare la nozione di trasporto parallelo di Levi-Civita. L'idea è di dare una prescrizione che permetta di dire se un vettore $A^{\mu}_{a}+dA^{\mu}_{a}$, posto in $x^{\alpha}_{a}+dx^{\alpha}_{a}$, è il vettore A^{μ}, posto in x^{α}_{a} e spostato parallelamente a se stesso di dx^{α}_{a}.

Se le x^{α}_{a} fossero coordinate cartesiane in $\mathbf{R}^n$, si avrebbe $\delta A^{\mu}_{a}=0$. In un sistema generico di coordinate b si potrebbe asserire solamente che

$$\delta A^{\mu}_{b}=-\Gamma^{\mu}_{\alpha\beta,\,b}A^{\alpha}_{b}\delta x^{\beta}_{b} \tag{6.6}$$

dove le $\Gamma^{\mu}_{\alpha\beta}$ sono certe funzioni del punto che dovranno essere specificate con argomenti plausibili. È possibile anzitutto calcolare le $\Gamma^{\mu}_{\alpha\beta,\,b}$ a partire da una carta cartesiana $\mathcal{U}_a$ in $\mathbf{R}^n$. Si ha

$$A^{\mu}_{b}=\frac{\partial x^{\mu}_{b}}{\partial x^{\lambda}_{a}}A^{\lambda}_{a}. \tag{6.7}$$

Spostando A^{μ}_{b} di δx^{α}_{b}, si ha, in $\mathcal{U}_a$, $\delta A^{\lambda}_{a}=0$ e quindi

$$\delta A^{\mu}_{b}=\frac{\partial^2 x^{\mu}_{b}}{\partial x^{\lambda}_{a}\partial x^{\beta}_{a}}\frac{\partial x^{\beta}_{a}}{\partial x^{\alpha}_{b}}\frac{\partial x^{\lambda}_{a}}{\partial x^{\sigma}_{b}}A^{\sigma}_{b}\delta x^{\alpha}_{b}; \tag{6.8}$$

pertanto, per l'osservazione 2,

$$\Gamma^{\mu}_{\alpha\sigma,\,b}=-\frac{\partial^2 x^{\mu}_{b}}{\partial x^{\lambda}_{a}\partial x^{\beta}_{a}}\frac{\partial x^{\beta}_{a}}{\partial x^{\alpha}_{b}}\frac{\partial x^{\lambda}_{a}}{\partial x^{\sigma}_{b}}=\frac{\partial x^{\mu}_{b}}{\partial x^{\lambda}_{a}}\frac{\partial^2 x^{\lambda}_{a}}{\partial x^{\sigma}_{b}\partial x^{\alpha}_{b}} \tag{6.9}$$

e anche:

$$\Gamma^{\mu}_{\alpha\sigma,\,b}=\Gamma^{\mu}_{\sigma\alpha,\,b}. \tag{6.10}$$

Inoltre, si può imporre che la lunghezza di A^{μ}, o meglio $A^2=A^{\mu}_{a}A^{\nu}_{a}\eta_{\mu\nu}$, rimanga la stessa durante il trasporto parallelo. Usando la (6.7) e la (6.3) applicata a $g_{\mu\nu,\,a}=\eta_{\mu\nu}$, si ha $A^2=g_{\alpha\beta,\,b}(x)A^{\alpha}_{b}A^{\beta}_{b}$; la condizione diventa

$$\delta(A^2)=\left(\frac{\partial g_{\alpha\beta,\,b}}{\partial x^{\mu}_{b}}-g_{\alpha\nu,\,b}\Gamma^{\nu}_{\mu\beta}-g_{\nu\beta,\,b}\Gamma^{\nu}_{\mu\alpha}\right)A^{\alpha}_{b}A^{\beta}_{b}\delta x^{\mu}_{b}=0$$

e, dovendo questa essere verificata per qualunque A^{α}_{b} e δx^{μ}_{b}, conduce a

$$\frac{\partial g_{\alpha\beta,\,b}}{\partial x^{\lambda}_{b}}=\Gamma^{\sigma}_{\alpha\lambda,\,b}g_{\sigma\beta,\,b}+\Gamma^{\sigma}_{\beta\lambda,\,b}g_{\alpha\sigma,\,b}. \tag{6.11}$$

Permutando circolarmente gli indici α, β, λ, si ottiene

$$\Gamma^{\sigma}_{\alpha\lambda}=\frac{1}{2}g^{\sigma\beta}\left(\frac{\partial g_{\alpha\beta}}{\partial x^{\lambda}}+\frac{\partial g_{\lambda\beta}}{\partial x^{\alpha}}-\frac{\partial g_{\alpha\lambda}}{\partial x^{\beta}}\right) \tag{6.12}$$

che risolve completamente il problema di definire le $\Gamma^{\sigma}_{\alpha\lambda}$, e quindi il trasporto parallelo in uno spazio-tempo qualsiasi. Le $\Gamma^{\sigma}_{\alpha\lambda}$ sono chiamati simboli di Christoffel di seconda specie, oppure coefficienti di connessione della varietà di Riemann. Per quanto detto, essi non sono tensori: se fossero tensori, il loro annullarsi in un $\mathbf{R}^n$ di una carta $(\mathcal{U}_a, \mathcal{C}_a)$ cartesiana, li renderebbe nulli sempre, in contraddizione con la (6.8).

Le proprietà di trasformazione di $\Gamma^{\sigma}_{\alpha\lambda}$ si deducono dalla (6.11):

$$\Gamma^{\sigma}_{\alpha\lambda,\,b}=\Gamma^{\tau}_{\mu\rho,\,a}\frac{\partial x^{\sigma}_{b}}{\partial x^{\tau}_{a}}\frac{\partial x^{\rho}_{a}}{\partial x^{\lambda}_{b}}\frac{\partial x^{\mu}_{a}}{\partial x^{\alpha}_{b}}+\frac{\partial^2 x^{\mu}_{a}}{\partial x^{\alpha}_{b}\partial x^{\lambda}_{b}}\frac{\partial x^{\sigma}_{b}}{\partial x^{\mu}_{a}}; \tag{6.13}$$

questa generalizza la (6.9), dato che $\Gamma^{\tau}_{\mu\rho,\,a}=0$ se $(\mathcal{U}_a, \mathcal{C}_a)$ è cartesiana. Come si vede, la (6.13) consiste di due termini, il primo dei quali, da solo, indicherebbe una struttura tensoriale in $\Gamma^{\sigma}_{\alpha\lambda}$.

6.3. Derivazione covariante.

Ora si vuol calcolare la variazione *effettiva* di un vettore, cioè si cerca di eliminare l'effetto della curvatura delle coordinate. Per raggiungere lo scopo, verrà calcolato il coefficiente differenziale della variazione del vettore nel modo seguente. Siano x^{α}, $x^{\alpha}+dx^{\alpha}$ le coordinate di due punti vicini P_1, P_2 in una carta e sia $A^{\mu}(x)$ un campo vettoriale. Si calcola la variazione di $A^{\mu}(x)$ in due tempi: anzitutto, si sposta il vettore $A^{\mu}(x)$ dal punto P_1 nel punto P_2 usando la (6.6): $A^{\mu}(x)+\delta A^{\mu}=A^{\mu}(x)-\Gamma^{\mu}_{\alpha\beta}A^{\alpha}dx^{\beta}$; e poi si forma la differenza

$$A^{\mu}{}_{,\beta}dx^{\beta}\equiv A^{\mu}(x+dx)-(A^{\mu}(x)+\delta A^{\mu})=\left(\frac{\partial A^{\mu}}{\partial x^{\beta}}+\Gamma^{\mu}_{\alpha\beta}A^{\alpha}\right)dx^{\beta}, \tag{6.14}$$

che si chiama incremento effettivo di A^{μ}; si chiamerà inoltre derivata covariante di A la quantità

$$A^{\mu}{}_{,\beta}=\frac{\partial A^{\mu}}{\partial x^{\beta}}+\Gamma^{\mu}_{\alpha\beta}A^{\alpha}. \tag{6.15}$$

Questa definizione è ovviamente giustificata dal fatto che, dalle (6.3) e (6.13), segue

$$\left(\frac{\partial A^{\mu}}{\partial x^{\beta}}+\Gamma^{\mu}_{\alpha\beta}A^{\alpha}\right)_a=\left(\frac{\partial A^{\lambda}}{\partial x^{\sigma}}+\Gamma^{\lambda}_{\rho\sigma}A^{\rho}\right)_b\frac{\partial x^{\mu}_{a}}{\partial x^{\lambda}_{b}}\frac{\partial x^{\sigma}_{b}}{\partial x^{\beta}_{a}} \tag{6.16}$$

che è la legge di trasformazione di un tensore di tipo (1, 1). Analogamente si può provare che, se A_{μ} è un vettore covariante, la derivata assoluta

$$A_{\mu,\beta}=\frac{\partial A_{\mu}}{\partial x^{\beta}}-\Gamma^{\nu}_{\mu\beta}A_{\nu} \tag{6.17}$$

è pure un tensore, di tipo (0, 2).

Generalizzando questi risultati, la regola 5 del § 3.2 deve essere sostituita con la seguente regola:

REGOLA 5. Se $T^{\mu_1\ldots\mu_q}_{\nu_1\ldots\nu_p}$ è un tensore di tipo (q,p), la derivata covariante data da

$$(6.18)\qquad T^{\mu_1\ldots\mu_q}_{\nu_1\ldots\nu_p,\beta}=\frac{\partial T^{\mu_1\ldots\mu_q}_{\nu_1\ldots\nu_p}}{\partial x^\beta}+\Gamma^{\mu_1}_{\alpha\beta}T^{\alpha,\mu_2\ldots\mu_q}_{\nu_1\ldots\nu_p}+\Gamma^{\mu_2}_{\alpha\beta}T^{\mu_1\alpha\mu_3\ldots\mu_q}_{\nu_1\ldots\nu_p}+\ldots+$$
$$+\Gamma^{\mu_q}_{\alpha\beta}T^{\mu_1\ldots\mu_{q-1}\alpha}_{\nu_1\ldots\nu_p}-\Gamma^{\alpha}_{\nu_1\beta}T^{\mu_1\ldots\mu_q}_{\alpha\nu_2\ldots\nu_p}-\Gamma^{\alpha}_{\nu_2\beta}T^{\mu_1\ldots\mu_q}_{\nu_1\alpha\nu_3\ldots\nu_p}-\ldots-$$
$$-\Gamma^{\alpha}_{\nu_p\beta}T^{\mu_1\ldots\mu_q}_{\nu_1\ldots\nu_{p-1}\alpha}$$

è un tensore di tipo $(q,p+1)$.

Verrà ora applicata la (6.18) ad alcuni esempi rilevanti. Si vede facilmente che risulta

$$(6.19)\qquad g_{\mu\nu,\lambda}=\frac{\partial g_{\mu\nu}}{\partial x^\lambda}-\Gamma^{\alpha}_{\mu\lambda}g_{\alpha\nu}-\Gamma^{\alpha}_{\lambda\nu}g_{\mu\alpha}=0,$$

per cui il tensore metrico è covariantemente costante. In modo analogo

$$(6.20)\qquad \delta^\mu_{\nu,\beta}=\frac{\partial\delta^\mu_\nu}{\partial x^\beta}+\Gamma^{\mu}_{\alpha\beta}\delta^\alpha_\nu-\Gamma^{\alpha}_{\nu\beta}\delta^\mu_\alpha=0.$$

Dalla (6.18) si ricava anche la regola di Leibniz per la derivata di prodotti tensoriali; ad esempio:

$$(6.21)\qquad (A_\mu B^\lambda)_{,\beta}=A_{\mu,\beta}B^\lambda+A_\mu B^\lambda{}_{,\beta}$$

$$(6.22)\qquad (\Phi A_\mu)_{,\beta}=\Phi_{,\beta}A_\mu+\Phi A_{\mu,\beta}.$$

Cosí, applicando (6.21) all'identità

$$(6.23)\qquad g^{\mu\nu}g_{\nu\lambda}=\delta^\mu_\lambda$$

e usando le (6.19) e (6.20), si trova

$$(g^{\mu\nu},g_{\nu\lambda})_{,\beta}=g^{\mu\nu}{}_{,\beta}g_{\nu\lambda}+g^{\mu\nu}g_{\nu\lambda,\beta}=\delta^\mu_{\lambda,\beta}=0$$

e, contraendo con $g^{\lambda\alpha}$,

$$(6.24)\qquad g^{\mu\alpha}{}_{,\beta}=\frac{\partial g^{\mu\alpha}}{\partial x^\beta}+\Gamma^{\mu}_{\sigma\beta}g^{\sigma\alpha}+\Gamma^{\alpha}_{\sigma\beta}g^{\mu\sigma}=0.$$

Sia ancora $|g_a|=\det g_a$. Dalla (6.4) si ricava $|g_a|=(\det\mathcal{J}_{ba})^2|g_b|$. Dunque il segno di $|g_a|$ non dipende dalla carta a; in $\mathfrak{M}$ $|g_a|$ è negativo, dato che in una carta localmente inerziale deve essere $g_{\mu\nu,a}=\eta_{\mu\nu}$ e $\det\eta=-1$. Si lascerà sempre indeterminata la dimensione n e si supporrà $|g_a|<0$ in tutti i casi che verranno esaminati (a meno di non stabilire esplicitamente il contrario). Conviene introdurre la quantità $\pm\sqrt{-|g_a|}$, fissando il segno relativo della $\sqrt{\ }$ tra carte diverse a,b, tramite la

$$(6.25)\qquad \sqrt{-|g_a|}=(\det\mathcal{J}_{ba})\sqrt{-|g_b|}.$$

Questa relazione si può interpretare dicendo che esprime l'invarianza dell'elemento di volume:

$$(6.26)\qquad dV^n\equiv\sqrt{-|g_a|}\,dx^1_a dx^2_a\ldots dx^n_a=\sqrt{-|g_b|}\,dx^1_b dx^2_b\ldots dx^n_b.$$

Si definirà quindi il volume di un simplesso

$$(6.27)\qquad V(\sigma_n)=\int_{\sigma_n}dV^n$$

e quello di una regione qualsiasi triangolabile Ω:

$$(6.28)\qquad V(\Omega)=\int_\Omega dV^n.$$

Infine sia dato

$$(6.29.1)\qquad \varepsilon_{(\nu),a}\equiv\varepsilon_{\nu_1\ldots\nu_n,a}=\sqrt{-|g_a|}(-1)^{p(\nu)}$$

se tutte le ν_i sono differenti e formano una permutazione di parità $p(\nu)$ dei numeri $1,\ldots,n$, e sia

$$(6.29.2)\qquad \varepsilon_{\nu_1\ldots\nu_n,a}=0\qquad\text{se}\quad \nu_h=\nu_k,$$

cioè se almeno due degli indici ν_i coincidono. È facile verificare che $\varepsilon_{(\nu),a}$ è un tensore di tipo $(0,n)$ e cioè che le (6.29), assegnate nella carta a, valgono automaticamente in tutte le carte per via della (6.3). Esso generalizza il tensore di Levi-Civita. La derivata covariante di $\varepsilon_{(\nu)}$ risulta essere nulla. Secondo l'osservazione 4, si possono alzare successivamente tutti gli indici di $\varepsilon_{(\nu)}$, ottenendo $\varepsilon^{(\mu)}=(\sqrt{-|g|})^{-1}(-1)^{p(\mu)}$, se (μ) è una permutazione di $1,\ldots,n$, $\varepsilon^{(\mu)}=0$, se vi sono almeno due indici coincidenti. Usando la (6.29), è possibile presentare la (6.26) in forma esplicitamente *scalare*:

$$(6.30)\qquad dV^n=\frac{1}{n!}\varepsilon_{\nu_1\nu_2\ldots\nu_n}dx^{\nu_1}dx^{\nu_2}\ldots dx^{\nu_n}.$$

Si chiama densità scalare una quantità ρ_a se

$$(6.31)\qquad \rho_a=\det\mathcal{J}_{ba}\rho_b,$$

che generalizza la (6.25). $\sqrt{-|g_a|}$ è dunque una densità; si noti allora che da ρ_a è sempre possibile formare lo scalare $\rho_a\sqrt{-|g_a|}$. Inoltre la densità scalare ρ_a origina anche il tensore

$$(6.32)\qquad \rho_{(\nu)}=\rho_a\sqrt{-|g_a|}\,\varepsilon_{(\nu)}.$$

6.4. La curvatura e il tensore di Riemann.

Normalmente, per una funzione qualsiasi $f(x,y)$ si ha $\partial^2 f/\partial x\partial y=\partial^2 f/\partial y\partial x$. In generale ciò significa che $f_{,\alpha\beta}=f_{,\beta\alpha}$, dove α, β sono indici di derivazione covariante e f è uno scalare. Infatti, dalla (6.18),

$$f_{,\alpha\beta}=\frac{\partial^2 f}{\partial x^\alpha\partial x^\beta}-\Gamma^{\lambda}_{\alpha\beta}\frac{\partial f}{\partial x^\lambda}=f_{,\beta\alpha}.$$

Questa simmetria non è valida per un tensore generico. Sia ad esempio A_λ un vettore; dalla regola 5 si avrà

$$A_{\lambda,\,\alpha\beta}=\frac{\partial^2 A_\lambda}{\partial x^\alpha \partial x^\beta}-\frac{\partial \Gamma^\nu_{\lambda\alpha}}{\partial x^\beta}A_\nu-\Gamma^\nu_{\lambda\alpha}\frac{\partial A_\nu}{\partial x^\beta}-\Gamma^\sigma_{\alpha\beta}\left(\frac{\partial A_\lambda}{\partial x^\sigma}-\Gamma^\nu_{\lambda\sigma}A_\nu\right)-\Gamma^\sigma_{\lambda\beta}\left(\frac{\partial A_\sigma}{\partial x^\alpha}-\Gamma^\nu_{\sigma\alpha}A_\nu\right),$$

per cui

$$A_{\lambda,\,\alpha\beta}-A_{\lambda,\,\beta\alpha}=\left(-\frac{\partial \Gamma^\nu_{\lambda\alpha}}{\partial x^\beta}+\frac{\partial \Gamma^\nu_{\lambda\beta}}{\partial x^\alpha}+\Gamma^\sigma_{\lambda\beta}\Gamma^\nu_{\sigma\alpha}-\Gamma^\sigma_{\lambda\alpha}\Gamma^\nu_{\sigma\beta}\right)A_\nu\equiv R_\lambda{}^\nu{}_{\alpha\beta}A_\nu, \tag{6.33}$$

avendo posto

$$R_\lambda{}^\nu{}_{\alpha\beta}=-\frac{\partial \Gamma^\nu_{\lambda\alpha}}{\partial x^\beta}+\frac{\partial \Gamma^\nu_{\lambda\beta}}{\partial x^\alpha}+\Gamma^\sigma_{\lambda\beta}\Gamma^\nu_{\sigma\alpha}-\Gamma^\sigma_{\lambda\alpha}\Gamma^\nu_{\sigma\beta}; \tag{6.34}$$

le quantità $R_\lambda{}^\nu{}_{\alpha\beta}$ sono le componenti del cosiddetto tensore di Riemann. Esse devono costituire un tensore in quanto la (6.33) è una formula tensoriale valida per qualunque vettore A_ν. La dimostrazione diretta di questo fatto, a partire dalla (6.34), risulta molto complicata. Accanto a $R_\lambda{}^\nu{}_{\alpha\beta}$, verranno usati anche il tensore contratto $R_{\lambda\alpha}=R_\lambda{}^\nu{}_{\alpha\nu}$ e la curvatura scalare $R=g^{\lambda\alpha}R_{\lambda\alpha}$.

Per comprendere il significato geometrico del tensore di Riemann, si consideri $P\in\mathfrak{M}$ e sia $A^\mu(P)$ un vettore inizialmente definito in P, che s'intende trasportare parallelamente a se stesso lungo una curva chiusa γ parametrizzata da $x^\alpha(s)$, $0\leq s\leq 1$, infinitesima, con $x^\alpha_P=x^\alpha(0)=x^\alpha(1)$. Sia $A^\mu(1)$ il vettore ottenuto da $A^\mu(0)$ per trasporto parallelo lungo γ. Si ha, secondo la (6.6),

$$dA^\mu(s)=-\Gamma^\mu_{\alpha\beta}(s)A^\alpha(s)\,dx^\beta, \tag{6.35}$$

dove tutte le quantità sono calcolate in s, da cui

$$A^\mu(1)-A^\mu(0)\equiv\Delta A^\mu=-\int_{s=0}^{s=1}\Gamma^\mu_{\alpha\beta}(s)A^\alpha(s)\,dx^\beta.$$

Posto $\xi^\beta(s)=x^\beta(s)-x^\beta(0)$, sviluppando $\Gamma^\mu_{\alpha\beta}(s)$ e $A^\alpha(s)$ al primo ordine in ξ^β e trascurando il termine in $(\xi)^2 d\xi$, si trova

$$\Delta A^\mu\simeq\left(-\left.\frac{\partial \Gamma^\mu_{\sigma\beta}}{\partial x^\nu}\right|_{s=0}+\Gamma^\mu_{\alpha\beta}(0)\Gamma^\alpha_{\sigma\nu}(0)\right)A^\sigma(0)\int_\gamma \xi^\nu\,d\xi^\beta.$$

Poiché γ è chiusa, si ha

$$\int_\gamma(\xi^\nu d\xi^\beta+\xi^\beta d\xi^\nu)=\int_\gamma d(\xi^\beta\xi^\nu)=0,$$

per cui

$$\Delta A^\mu\simeq-\frac{1}{2}\int_\gamma(\xi^\nu d\xi^\beta-\xi^\beta d\xi^\nu)\cdot R_\sigma{}^\mu{}_{\nu\beta}(0)A^\sigma(0),$$

vera al secondo ordine in ξ. La quantità $\int_\gamma(\xi^\beta d\xi^\nu-\xi^\nu d\xi^\beta)/2$ individua l'elemento d'area racchiuso da γ in grandezza e orientazione; dunque, il vettore A^μ, trasportato parallelamente lungo γ, cambia di un ammontare proporzionale: al vettore A^μ, all'area racchiusa da γ e alla curvatura $R_\sigma{}^\mu{}_{\nu\beta}$. Ciò generalizza quanto

visto al § 5.4 a piú dimensioni; le $R_\sigma{}^\mu{}_{\nu\beta}$ devono essere associate all'angolo della curvatura gaussiana, secondo le idee originali di Riemann.

6.5. Coordinate inerziali.

Si è ora in grado di discutere in dettaglio l'esistenza o meno di una carta locale $(\mathcal{U}_a, \Phi_a)$ tale che in $P\in\mathcal{U}_a$ si abbia $g_{\mu\nu,\,a}(P)=\eta_{\mu\nu}$. È noto che

$$g_{\mu\nu,\,a}=g_{\alpha\beta,\,b}\frac{\partial x^\alpha_b}{\partial x^\mu_a}\frac{\partial x^\beta_b}{\partial x^\nu_a},$$

dove le $\partial x^\alpha_b/\partial x^\mu_a$ possono essere scelte arbitrariamente in P, ma non nella $\mathcal{U}_a$ completa; ragionando (per fissare le idee) nel caso $n=4$, queste derivate contano come sedici parametri, dieci dei quali possono essere fissati in modo da avere $g_{\mu\nu,\,a}=\eta_{\mu\nu}$; quanto ai sei rimanenti, essi corrispondono alle trasformazioni di Lorentz che, come è stato ampiamente discusso, non alterano $\eta_{\mu\nu}$. Supponendo quindi *ab initio* $g_{\mu\nu,\,a}|_P=\eta_{\mu\nu}$ e restringendosi inoltre a una classe di carte $\mathcal{A}_1(P)$ tale che, se $b\in\mathcal{A}_1(P)$, sia $\partial x^\mu_b/\partial x^\nu_a|_P=\delta^\mu_\nu$, cosí che $a\in\mathcal{A}_1(P)$, si ha allora $g_{\mu\nu,\,b}|_P=\eta_{\mu\nu}$ e la (6.13) diventa

$$\Gamma^\sigma_{\alpha\lambda,\,b}|_P=\left(\Gamma^\sigma_{\alpha\lambda,\,a}+\frac{\partial^2 x^\sigma_a}{\partial x^\alpha_b\,\partial x^\lambda_b}\right)_P,$$

da intendersi ristretta al punto P, con $a,b\in\mathcal{A}_1(P)$. In essa compaiono $n^2(n+1)/2$ derivate seconde e componenti tensoriali distinte (40 se $n=4$), e le derivate seconde sono tuttora arbitrarie in P; pertanto, se si pone $\Gamma^\sigma_{\alpha\lambda,\,a}=0$, si possono sempre scegliere le derivate seconde in modo da avere $\Gamma^\sigma_{\alpha\lambda,\,b}|_P=0$. Ottenuto questo risultato, la carta b si dirà *inerziale*, e in essa tutte le derivate covarianti si riducono in P a derivate ordinarie. Si noti tuttavia che, anche in P, in generale $\partial\Gamma^\sigma_{\alpha\lambda}/\partial x^\beta\neq 0$; e, se in P è $\Gamma^\sigma_{\alpha\lambda}=0$, la (6.11) dà subito $\partial g_{\alpha\lambda}/\partial x^\sigma=0$, $\forall\alpha,\lambda,\sigma$.

L'esistenza di una carta inerziale in P realizza dunque il principio di equivalenza, in quanto tutte le derivate si riducono alle stesse operazioni della relatività ristretta e la stazionarietà della metrica garantisce che $g_{\alpha\lambda}=\eta_{\alpha\lambda}$ approssimativamente in un intorno sufficientemente piccolo di P.

Ci si può dunque limitare, *ab initio*, ad esaminare una nuova classe di carte $\mathcal{A}_2(P)$, tale che $\Gamma^\sigma_{\alpha\lambda,\,b}=0$, $\forall b\in\mathcal{A}_2(P)$; pertanto, se $a,b\in\mathcal{A}_2(P)$, si avrà in P:

$$\left.\frac{\partial x^\mu_a}{\partial x^\lambda_b}\right|_P=\delta^\mu_\lambda \qquad \left.\frac{\partial^2 x^\mu_a}{\partial x^\lambda_b\,\partial x^\nu_b}\right|_P=0. \tag{6.36}$$

Usando queste restrizioni, la derivata della (6.13) rispetto ad un generico x^ν_b, effettuata nel punto P, diventa:

$$\left.\frac{\partial\Gamma^\sigma_{\alpha\lambda,\,b}}{\partial x^\nu_b}\right|_P=\left.\frac{\partial\Gamma^\sigma_{\alpha\lambda,\,a}}{\partial x^\nu_a}\right|_P+\left.\frac{\partial^3 x^\sigma_a}{\partial x^\alpha_b\,\partial x^\lambda_b\,\partial x^\nu_b}\right|_P.$$

Qui compaiono $n^2(n+1)(n+2)/6$ derivate terze (80 se $n=4$) e $n^3(n+1)/2$ derivate $\partial\Gamma^\sigma_{\alpha\lambda}/\partial x^\nu|_P$ (160 se $n=4$); di queste ultime, solamente $(n(n+1)/2)^2$ sono

indipendenti, poiché tutte quante esprimibili in P a mezzo delle $\partial^2 g_{\alpha\beta}/\partial x^\mu \partial x^\nu|_P$ (100 se $n=4$). Pertanto non è possibile scegliere le derivate terze in modo da annullare tutte le $\partial \Gamma^\sigma_{\alpha\lambda,b}/\partial x^\nu_b|_P$; tuttavia si potranno esprimere le $\partial^2 g_{\alpha\beta}/\partial x^\mu \partial x^\nu|_P$ per mezzo di un numero limitato di esse: $(n(n+1)/2)^2-[n^2(n+1)(n+2)/6]=$ $=n^2(n^2-1)/12$, cioè 20 se $n=4$. Queste venti quantità sono legate al tensore di Riemann dalla

$$(6.37)\qquad R_\lambda{}^\nu{}_{\alpha\beta,a}|_P=\frac{1}{2}\eta^{\nu\sigma}\left(\frac{\partial^2 g_{\sigma\beta}}{\partial x^\lambda \partial x^\alpha}+\frac{\partial^2 g_{\lambda\alpha}}{\partial x^\sigma \partial x^\beta}-\frac{\partial^2 g_{\sigma\alpha}}{\partial x^\lambda \partial x^\beta}-\frac{\partial^2 g_{\lambda\beta}}{\partial x^\sigma \partial x^\alpha}\right)_{a,P},$$

da intendersi ristretta a P con $a\in \mathcal{A}_2(P)$. In modo piú simmetrico, introducendo il tensore

$$R_{\tau\lambda,\alpha\beta}\equiv g_{\nu\tau}R_\lambda{}^\nu{}_{\alpha\beta}$$

si ha:

$$(6.37')\qquad R_{\tau\lambda,\alpha\beta}|_{a,P}=\frac{1}{2}\left(\frac{\partial^2 g_{\tau\beta}}{\partial x^\lambda \partial x^\alpha}+\frac{\partial^2 g_{\lambda\alpha}}{\partial x^\tau \partial x^\beta}-\frac{\partial^2 g_{\tau\alpha}}{\partial x^\lambda \partial x^\beta}-\frac{\partial^2 g_{\lambda\beta}}{\partial x^\tau \partial x^\alpha}\right)_{a,P}.$$

Dalla (6.37′) si vede che il tensore $R_{\tau\lambda,\alpha\beta}$ è dotato delle simmetrie

$$(6.38.1)\qquad R_{\tau\lambda,\alpha\beta}=-R_{\tau\lambda,\beta\alpha}$$

$$(6.38.2)\qquad R_{\tau\lambda,\alpha\beta}=R_{\alpha\beta,\tau\lambda}$$

$$(6.38.3)\qquad R_{\tau\alpha,\beta\lambda}+R_{\tau\beta,\lambda\alpha}+R_{\tau\lambda,\alpha\beta}=0;$$

l'analisi combinatoria di queste simmetrie mostra che $R_{\tau\lambda,\alpha\beta}$ ha effettivamente solo $n^2(n^2-1)/12$ componenti in generale distinte. Difatti valgono le condizioni originalmente proposte da Riemann,

$$(6.39)\qquad \left.\frac{\partial^2 g_{\tau\beta}}{\partial x^\lambda \partial x^\alpha}\right|_P=-\frac{1}{3}(R_{\tau\lambda,\beta\alpha}+R_{\beta\lambda,\tau\alpha})_P,$$

che, insieme alle $g_{\mu\nu,a}=\eta_{\mu\nu}$ e $\Gamma^\sigma_{\mu\nu,a}=0$, definiscono il cosiddetto sistema di coordinate normali di Riemann nell'intorno di P. Ponendo $x^\alpha_a=0$ in P, si ha al second'ordine, nelle x^α_a,

$$(6.40)\qquad g_{\mu\nu,a}=\eta_{\mu\nu}-\frac{1}{3}R_{\mu\lambda,\nu\alpha}|_P\, x^\lambda_a x^\alpha_a+\ldots;$$

i termini di ordine superiore sembrano non rivestire al momento grande interesse.

Relativamente al tensore di Riemann, si riporteranno ancora le identità di Bianchi

$$(6.41)\qquad R_{\tau\lambda,\alpha\beta;\rho}+R_{\tau\lambda,\beta\rho;\alpha}+R_{\tau\lambda,\rho\alpha;\beta}=0,$$

la cui esistenza dipende dalla possibilità di scegliere infinite superfici diverse Σ aventi come bordo il contorno chiuso $\gamma=\partial\Sigma$ di cui al § 6.4. Contraendo la (6.41) con $g^{\tau\beta}$ e con $g^{\lambda\alpha}$ si ha

$$(6.42)\qquad R_{;\rho}-2R^\alpha{}_{\rho;\alpha}=0$$

ed introducendo il tensore

$$(6.43)\qquad G^\alpha{}_\rho=R^\alpha{}_\rho-\frac{1}{2}\delta^\alpha_\rho R$$

si può riscrivere la (6.42) nella forma

$$(6.44)\qquad G^\alpha{}_{\rho;\alpha}=0$$

alla quale si ritornerà trattando delle equazioni di campo einsteiniane.

6.6. Geodetiche. Riferimenti gaussiani.

Una linea γ nella $\mathfrak{M}^n$ si dice di tipo tempo se tra due punti vicini x^α, $x^\alpha+dx^\alpha$ di γ vale la disuguaglianza $ds^2=g_{\mu\nu}dx^\mu dx^\nu<0$. In questo caso l'integrale

$$(6.45)\qquad \int_1^2 -\sqrt{g_{\mu\nu}dx^\mu dx^\nu}=\int_1^2 du=u_2-u_1$$

dicesi intervallo di tempo proprio tra i punti 1 e 2 della curva. γ sarà specificata esplicitamente assegnando $x^\alpha=x^\alpha(s)$ in funzione di un parametro s che cresce monotonamente lungo la direzione positiva del tempo. Questo parametro può essere la u stessa definita esplicitamente dall'integrale (6.45). In tal caso:

$$(6.46)\qquad g_{\mu\nu}\frac{dx^\mu}{du}\frac{dx^\nu}{du}=-1.$$

Dati due punti qualsiasi P_1, P_2 in $\mathfrak{M}^n$, sorge il problema di trovare l'arco di curva γ congiungente P_1 a P_2 per il quale l'integrale (6.45) risulti massimo (se la metrica fosse definita positiva, si richiederebbe il minimo). Una condizione necessaria (ma, per archi abbastanza lunghi, non sufficiente) affinché ciò si verifichi, è che $\int_{P_1}^{P_2} du$ sia stazionario rispetto a variazioni infinitesime delle $x^\alpha(u)$. Tale condizione dà

$$(6.47)\qquad \frac{d^2x^\alpha}{du^2}+\Gamma^\alpha_{\lambda\nu}\frac{dx^\lambda}{du}\frac{dx^\nu}{du}=0,$$

che è la notissima equazione per le geodetiche; essa svolge un ruolo fondamentale nella meccanica del moto di un corpo puntiforme in un campo gravitazionale. Si noti che il vettore tangente $T^\alpha=dx^\alpha/du$ viene trasportato parallelamente a se stesso lungo una geodetica; infatti, ponendo $\delta T^\alpha=(dT^\alpha/du)\delta u$, si ha $\delta T^\alpha=-\Gamma^\alpha_{\lambda\nu}T^\nu\delta x^\lambda$, che coincide con la (6.35).

Si considerino una varietà differenziabile $\mathfrak{M}^n$ dotata di metrica $g_{\mu\nu}$ definita positiva, e una ipersuperficie Σ^{n-1}. Sia $(\mathcal{U},\Phi)$ una carta di Σ^{n-1}; si consideri la geodetica γ_x in $\mathfrak{M}^n$ passante per $x\in\mathcal{U}$ ed ortogonale a Σ^{n-1}. Un punto generico della geodetica sarà identificato dalla coppia (t,x), dove t è l'arco di geodetica tra (t,x) e $(0,x)$. Questa coppia identifica una carta di $\mathfrak{M}^n$, purché t sia sufficientemente piccolo, altrimenti possono esservi piú geodetiche tra (t,x) e $(0,x)$. Euristicamente, esisterà in questa carta una metrica del tipo

$$(6.48)\qquad ds^2=g_{00}dt^2+2g_{0i}dt\,dx^i+g_{ik}dx^i dx^k.$$

Per quanto detto, una geodetica particolare di (6.48) è data da $x^i = \text{cost}$. Si ha dunque $ds = \sqrt{g_{00}}\, dt$; dalla definizione di t, si ha $ds = dt$, da cui $g_{00} = 1$. Inoltre, l'equazione delle geodetiche si scrive

$$0 = \frac{d^2x^\nu}{ds^2} = \frac{d^2x^\nu}{dt^2} = -\Gamma^\nu_{\alpha\beta}\frac{dx^\alpha}{dt}\frac{dx^\beta}{dt} = -\Gamma^\nu_{00},$$

che è identicamente soddisfatta purché $\Gamma^\nu_{00} = 0$, ossia $\partial g_{0\nu}/\partial t = 0$. $g_{0\nu}$ non dipende da t; d'altra parte, per $t = 0$, l'ortogonalità di γ_x a Σ^{n-1} implica $g_{0\nu} = 0$, quindi si ha sempre $g_{0\nu} = 0$. La metrica assume pertanto la forma semplificata gaussiana

$$ds^2 = dt^2 + g_{ik}dx^i dx^k. \tag{6.49}$$

Si può banalmente ripetere il ragionamento per metriche di tipo Minkowski, ove t sia una variabile temporale; si ottiene

$$ds^2 = -dt^2 + g_{ik}dx^i dx^k, \tag{6.50}$$

che viene impiegata in cosmologia. Una carta gaussiana è di necessità locale, diventando singolare lungo le caustiche focali della congruenza normale di geodetiche γ_x. Le geodetiche sono normali a tutte le ipersuperfici $t = \text{cost}$.

7. *Il campo gravitazionale.*

7.1. Equazioni del campo gravitazionale.

Nella relatività ristretta il moto di un corpo libero può essere derivato dal principio d'azione:

$$\delta\left(-mc\int_1^2 du\right) = 0 \tag{7.1}$$

dove u è il tempo proprio. Le equazioni che si ottengono sono: $d^2x^\alpha/du^2 = 0$. Secondo il principio di equivalenza, le stesse equazioni devono valere in un sistema localmente inerziale. Ma l'equazione delle geodetiche

$$\frac{d^2x^\alpha}{du^2} + \Gamma^\alpha_{\mu\nu}\frac{dx^\mu}{du}\frac{dx^\nu}{du} = 0 \tag{7.2}$$

si riduce alla $d^2x^\alpha/du^2 = 0$ quando $\Gamma^\alpha_{\mu\nu} = 0$ e ha carattere tensoriale, per cui si debbono assumere le (7.2) come equazioni del moto derivate dal principio variazionale con azione

$$S = -mc\int_1^2 du \tag{7.3}$$

dove du è lo stesso di (6.45).

Ci si ponga ora il problema di confrontare le (7.2) con le usuali equazioni del moto newtoniano. A questo scopo conviene sottolineare che, nonostante le $g_{\mu\nu}$ descrivano uno spazio curvo, la loro deviazione dalle $\eta_{\mu\nu}$ deve essere piccolissima, altrimenti si sarebbero già notate da tempo, entro un campo gravitazionale, vistose deviazioni dalla geometria euclidea. Quindi si può porre $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, dove $h_{\mu\nu} \ll 1$. Inoltre si ha a che fare (ad esempio nel campo terrestre) con campi quasi statici, cioè che variano molto lentamente nel tempo, per cui conviene porre $\partial h_{\mu\nu}/\partial x^0 = 0$. Infine il moto dei gravi è di norma non relativistico, per cui si trascureranno dx^i/du rispetto a $dx^0/du \simeq 1$. Trascurando i termini quadratici in $h_{\mu\nu}$, segue approssimativamente

$$\frac{d^2x^\alpha}{du^2} + \Gamma^\alpha_{00}\left(\frac{dx^0}{du}\right)^2 = 0. \tag{7.4}$$

La (7.4), con $\alpha = 0$, si riduce a $d^2x^0/du^2 = 0$, ossia $x^0 = au + b$; non c'è perdita di generalità nel supporre $a = 1$, $b = 0$, $x^0 = u$. Inoltre

$$\frac{d^2x^i}{du^2} = -\Gamma^i_{00} = \frac{1}{2}\frac{\partial h_{00}}{\partial x^i}. \tag{7.5}$$

Ricordando che le equazioni di Newton si scrivono

$$\frac{d^2x^i}{dt^2} = -\frac{\partial\Phi}{\partial x^i}, \tag{7.6}$$

dove Φ è il potenziale gravitazionale, essendo $u = ct$, si vede che la (7.6) s'identifica con la (7.5) se si pone $-(2\Phi/c^2) = h_{00}$ e quindi

$$g_{00} = -\left(1 + \frac{2\Phi}{c^2}\right). \tag{7.7}$$

Si noti che, se m è la massa del grave, $m\Phi$ è l'energia gravitazionale e $2m\Phi/mc^2$ il rapporto tra energia gravitazionale ed equivalente energetico del corpo, che di solito è enormemente piú grande. Dunque $2\Phi/c^2 \ll 1$, come richiesto. Ad esempio, sulla Terra, $\Phi = -(GM/R)$, dove $M = 5{,}977 \cdot 10^{27}$ g e $R = 6{,}37 \cdot 10^8$ cm sono la massa e il raggio della Terra, G la costante di Newton: $G = 6{,}67 \cdot 10^{-8}$ g^{-1} cm^3 sec^{-2}. Si ha all'incirca $\Phi \simeq -6{,}24 \cdot 10^{11}$ cm^2 sec^{-2}, $h_{00} \simeq 1{,}39 \cdot 10^{-9}$. Invece, per il Sole, $M \simeq 10^{33}$ g, $R \simeq 0{,}75 \simeq 10^{11}$ cm, e $h_{00} \sim 10^{-6}$.

Si può scrivere

$$h_{00} = \frac{2GM}{Rc^2} = \left(\frac{2GM}{c^2}\right)\frac{1}{R} = \frac{R(M)}{R},$$

dove $R(M) = 2MG/c^2$ è il cosiddetto raggio di Schwarzschild della massa M. Le deviazioni visibili dalla geometria solita sorgono quando $R \sim R(M)$; per la Terra, $R(M)$ è dell'ordine del centimetro, per il Sole di un chilometro, per le stelle a neutroni $R(M)/R \sim 0{,}1$, e nei buchi neri, $R = R(M)$ per definizione.

Dalla discussione precedente segue che h_{00}, e quindi g_{00}, va interpretato come potenziale gravitazionale e Γ^i_{00} come accelerazione di gravità (a parte il fattore c^2). In generale le (7.2) esprimono l'influenza della gravitazione e dell'inerzia sui corpi. Non è possibile distinguere l'inerzia dalla gravitazione, in quanto nella (7.2) i due termini non hanno separatamente carattere tensoriale: cambiando carta, essi si trasformano l'uno nell'altro.

Fu una componente costante del pensiero di Einstein l'idea che un sistema fisico, ad esempio il campo gravitazionale, non potesse influenzarne un altro, ad esempio una particella, senza essere a sua volta influenzato. Occorre dunque un'equazione che leghi il campo gravitazionale $g_{\mu\nu}$ alla distribuzione della materia. La (7.7) simboleggia l'identità dell'ente geometrico $g_{\mu\nu}$ con il potenziale gravitazionale.

Nella gravitazione newtoniana la relazione tra Φ e la distribuzione di materia è data dall'equazione di Poisson

$$\Delta\Phi = 4\pi G\rho \tag{7.8}$$

dove ρ è la densità di materia. Conviene dunque supporre che la (7.8) sia una forma approssimata di altre equazioni relativistiche che legano la distribuzione di materia alle $g_{\mu\nu}$. Le equazioni esatte che rimpiazzano la (7.8) dovranno, per analogia, essere del secondo ordine nelle $g_{\mu\nu}$, lineari ed omogenee nelle derivate seconde, ed avere carattere tensoriale. Lo scalare Φ verrebbe quindi ad essere rimpiazzato da un tensore $g_{\mu\nu}$ che ha dieci componenti, per cui le equazioni del campo dovrebbero avere dieci componenti. Per questa ragione la descrizione della distribuzione di materia mediante la ρ è insoddisfacente, e la ρ deve a sua volta essere sostituita da un tensore $T_{\mu\nu}$, che contiene anche informazione sullo stato di moto dei corpi materiali.

Ci si attende dunque un'equazione del tipo

$$G_{\mu\nu} = KT_{\mu\nu} \tag{7.9}$$

dove le $G_{\mu\nu}$ sono componenti di un tensore formato con le derivate prime e seconde di $g_{\mu\nu}$ e lineare in queste. Si può dimostrare matematicamente che tutti i tensori di questa forma sono necessariamente del tipo

$$G_{\mu\nu} = R_{\mu\nu} + \lambda g_{\mu\nu} R, \tag{7.10}$$

dove $R_{\mu\nu}$ è il tensore contratto di Riemann (tensore di Ricci) e R è la curvatura scalare. Si scriverà la (7.9) come

$$R_{\mu\nu} + \lambda g_{\mu\nu} R = KT_{\mu\nu}. \tag{7.11}$$

In analogia con la meccanica newtoniana, le $T_{\mu\nu}$ devono obbedire a equazioni di continuità, che tensorialmente si scrivono

$$T^{\mu}{}_{\nu;\mu} = 0, \tag{7.12}$$

per cui si deve avere

$$R^{\mu}{}_{\nu;\mu} + \lambda R_{\nu} = 0, \tag{7.13}$$

equazione che si riduce alla (6.42) se $\lambda = -1/2$, condizione che si assumerà in quanto segue. Si ha dunque:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = kT_{\mu\nu} \tag{7.14}$$

come modello di teoria del campo gravitazionale, ove k è una costante da determinarsi. Si può anche supporre che la materia sia distribuita con densità ρ e

quadrivelocità dx^{α}/du, cosí che un'approssimazione sufficiente agli scopi qui perseguiti si ottiene ponendo

$$T_{\mu\nu} = \rho c^2 \frac{dx_{\mu}}{du}\frac{dx_{\nu}}{du}. \tag{7.15}$$

7.2. Limite statico delle equazioni di campo.

Si supponga ora $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, dove $h_{\mu\nu} \ll 1$. Trascurando i termini quadratici in $h_{\mu\nu}$, si trova:

$$R_{\tau\lambda\alpha\beta} = \frac{1}{2}\left(\frac{\partial^2 h_{\tau\beta}}{\partial x^{\lambda}\partial x^{\alpha}} + \frac{\partial^2 h_{\lambda\alpha}}{\partial x^{\tau}\partial x^{\beta}} - \frac{\partial^2 h_{\tau\alpha}}{\partial x^{\lambda}\partial x^{\beta}} - \frac{\partial^2 h_{\lambda\beta}}{\partial x^{\tau}\partial x^{\alpha}}\right). \tag{7.16}$$

Contraendo con $g^{\tau\beta}$ (ossia con $\eta^{\tau\beta}$ in questa approssimazione),

$$R_{\lambda\alpha} = \frac{1}{2}\left(\frac{\partial^2 h}{\partial x^{\lambda}\partial x^{\alpha}} + \Box h_{\lambda\alpha} - \frac{\partial h_{\alpha}}{\partial x^{\lambda}} - \frac{\partial h_{\lambda}}{\partial x^{\alpha}}\right) \tag{7.17}$$

dove

$$h = \eta^{\tau\beta} h_{\tau\beta} \qquad h_{\alpha} = \frac{\partial h_{\tau\alpha}}{\partial x^{\beta}}\eta^{\tau\beta}.$$

Infine

$$R = \Box h - \eta^{\lambda\alpha}\frac{\partial h_{\alpha}}{\partial x^{\lambda}}. \tag{7.18}$$

S'incontra ora un'arbitrarietà nel campo $h_{\mu\nu}$, che riflette il carattere tensoriale di $g_{\mu\nu}$. Operando con un cambiamento infinitesimo di coordinate del tipo $x^{\mu}_{b} = x^{\mu}_{a} + \xi^{\mu}(x)$, dove $\partial\xi^{\mu}/\partial x^{\lambda}_{b} \ll 1$, e trascurando i prodotti del tipo $h\xi$, ξ^2, ecc., si ottiene

$$h_{\mu\nu,a} = h_{\mu\nu,b} + \frac{\partial\xi_{\mu}}{\partial x^{\nu}_{a}} + \frac{\partial\xi_{\nu}}{\partial x^{\mu}_{a}},$$

cioè si ottiene un nuovo $h_{\mu\nu}$, con $h_{\mu\nu} \ll 1$, che offre in principio la stessa configurazione fisica di prima. Ne segue che si possono fissare le coordinate imponendo la cosiddetta condizione armonica sulle x^{μ}:

$$\Gamma^{\mu}_{\sigma\alpha} g^{\sigma\alpha} = 0. \tag{7.19}$$

Nell'approssimazione qui adottata essa implica

$$\frac{\partial h}{\partial x^{\mu}} = 2h_{\mu}, \tag{7.20}$$

per cui semplicemente

$$R_{\lambda\alpha} = \frac{1}{2}\Box h_{\lambda\alpha} \qquad R = \frac{1}{2}\Box h, \tag{7.21}$$

e la (7.14) prende la forma

$$\Box\left(h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} h\right) = 2kT_{\mu\nu}. \tag{7.22}$$

Contraendo con $\eta^{\mu\nu}$ e notando che $\delta^{\mu}_{\mu}=4$, si ottiene: $-\Box h=2kT$, dalla quale deriva la

$$\Box h_{\mu\nu}=2k\left(T_{\mu\nu}-\frac{1}{2}\eta_{\mu\nu}T\right). \tag{7.23}$$

Nel limite statico in cui $dx^0/du=1$, $dx^i/du=0$, si ha $T_{00}=\rho c^2$, $T_{ij}=T_{0i}=0$, ed $\eta_{\mu\nu}$ non dipende dal tempo. Pertanto $\Box h_{\mu\nu}=\Delta h_{\mu\nu}$ e, ponendo $\mu=\nu=0$, $\Delta h_{00}=2k[\rho-(\rho/2)]c^2=k\rho c^2$. Essendo $h_{00}=-(2\Phi/c^2)$, si trova $\Delta\Phi=-(c^4/2)k\rho$, che, confrontata con la (7.8), porta a porre

$$k=-\frac{8\pi G}{c^4}. \tag{7.24}$$

Si giunge dunque alle

$$R_{\mu\nu}-\frac{1}{2}g_{\mu\nu}R=-\frac{8\pi G}{c^4}T_{\mu\nu}, \tag{7.25}$$

che costituiscono le equazioni fondamentali del campo einsteiniano. Le R si misurano in cm^{-2}, le $T_{\mu\nu}/c^2$ in $\mathrm{g/cm}^2$; inoltre: $8\pi G/c^2=1{,}86\cdot 10^{-27}$ cm/g. Grosso modo, una densità di 1 $\mathrm{g/cm}^3$ corrisponde a $R\sim(4\cdot 10^{-14})^2\ \mathrm{cm}^{-2}$ (R generica componente della curvatura). Per vedere gli effetti della curvatura occorre avere una distribuzione di acqua con raggio $2{,}5\cdot 10^8$ km, poco piú del raggio dell'orbita terrestre.

Nel limite statico si ha anche

$$\Delta h_{ij}=-k\eta_{ij}T=\frac{8\pi G}{c^2}\rho\,\delta_{ij}, \tag{7.26}$$

per cui $h_{ij}=\delta_{ij}h_{00}$; appaiono dunque anche componenti spaziali di $h_{\mu\nu}$.

7.3. Lo spostamento gravitazionale verso il rosso.

Questo importante fenomeno ha fornito una delle basi osservative piú note della relatività generale. Tutti i corpi che salgono in un campo gravitazionale, rallentano e perdono energia cinetica. La perdita di energia della luce si manifesta invece nella diminuzione della frequenza, dato che $E=h\nu$, e non dalla velocità, che rimane (e deve rimanere) fissa sul valore c. L'argomento si può meglio precisare nel modo seguente.

Sia Φ un campo gravitazionale statico prodotto ad esempio da una stella. Siano A, B, due osservatori, il primo distante (alto) e il secondo vicino (basso) alla stella, di modo che $\Phi_A>\Phi_B$. Sia dato in A un orologio che scandisce il tempo con periodo T in una serie di eventi $t_n=nT$. A vede dunque passare il tempo proprio

$$T_A=T\sqrt{1+\frac{2\Phi_A}{c^2}}\simeq T\left(1+\frac{\Phi_A}{c^2}\right),$$

tra due battiti, dato che: $-du^2=g_{00}(dx_0)^2$, poiché gli elementi $(dx^i)^2$ sono nulli, essendo l'orologio fisso in A. Esiste inoltre una geodetica «messaggera» $x^i(t)$ tale che $x^i(0)$ sia in A e $x^i(t)$ sia in B. Ad ogni battito dell'orologio in A, un fotone parte da A e, viaggiando lungo tale geodetica, raggiunge B. B vede dunque gli arrivi succedersi ad intervalli

$$T_B\simeq T\left(1+\frac{\Phi_B}{c^2}\right).$$

Il rapporto

$$\frac{T_A}{T_B}\simeq\frac{1+\dfrac{\Phi_A}{c^2}}{1+\dfrac{\Phi_B}{c^2}}\simeq 1+\frac{\Phi_A-\Phi_B}{c^2}>1$$

dice che B percepisce i battiti accelerati rispetto a quanto accade ad A. Viceversa, A vede i fenomeni su B svolgersi a ritmo rallentato. In particolare, sia un atomo su B emettente radiazione con frequenza ν_B e periodo $1/\nu_B$; su A la radiazione giunge con frequenza

$$\nu_A\simeq\left(1+\frac{\Phi_B-\Phi_A}{c^2}\right)\nu_B<\nu_B, \tag{7.27}$$

che corrisponde al citato spostamento verso il rosso.

Il seguente esperimento concettuale illustra in modo diverso la necessità dello spostamento verso il rosso. Sia costruito un apparato formato da una colonna a fluido ascendente da B ad A, e da una colonna dove il fluido discende da A a B ricominciando il ciclo. Il fluido è composto da atomi aventi due livelli con differenza di energia $\Delta E=h\nu$. Se la massa di un atomo è m nel livello piú basso, varrà $m+\Delta E/c^2$ in quello piú alto. Durante il ciclo, in B gli atomi si diseccitano e la radiazione emessa con frequenza ν viene convogliata in A. Saliti su A con massa m, gli atomi ricevono la radiazione e vengono spostati sul livello eccitato, per poi discendere in B dove ricomincia il ciclo. Se non esistesse spostamento verso il rosso, l'apparato produrrebbe il moto perpetuo, dato che gli atomi discendono con peso piú grande. La differenza di lavoro per atomo sarebbe dunque:

$$(\Phi_A-\Phi_B)\times(\text{differenza di massa})=\frac{\Phi_A-\Phi_B}{c^2}\Delta E.$$

I fotoni in ascesa arrivano in A con frequenza

$$\nu_A=\left(1+\frac{\Phi_B-\Phi_A}{c^2}\right)\nu_B$$

e la loro perdita di energia è

$$h\nu_B-h\nu_A\simeq\frac{\Phi_A-\Phi_B}{c^2}h\nu_B=\frac{\Phi_A-\Phi_B}{c^2}\Delta E,$$

che compensa esattamente il guadagno precedente. L'effetto evita dunque la possibilità, ben piú disastrosa concettualmente, del moto perpetuo.

Ma a quanto ammonta in pratica l'effetto? Sulla Terra, per un dislivello di h metri, si ha $\Phi_A - \Phi_B = hg$, dove g è l'accelerazione di gravità. Dunque il cambiamento percentuale in frequenza

$$Z = \frac{\nu_B - \nu_A}{\nu_B} \sim \frac{\Phi_A - \Phi_B}{c^2} = \frac{hg}{c^2}$$

è piccolissimo: se il dislivello h è di 10 metri, $Z \approx 10^{-15}$. Tra il Sole e l'infinito si ha $Z \sim -\Phi/c^2 \sim 10^{-6}$. Nelle stelle a neutroni sale del 10 per cento e diventa infinito per i buchi neri.

L'effetto è stato rilevato sulla Terra da Rebka e Pound [1960] usando raggi γ di frequenza molto ben definita, prodotti dall'effetto Mossbauer, su un dislivello di poche decine di metri. Sul Sole l'effetto è presente, ma è mascherato dall'agitazione termica e dallo spostamento delle righe prodotto dai campi magnetici (effetto Zeeman). Esso è stato misurato sulle nane bianche, dove $Z \sim 10^{-4}$.

7.4. Le onde gravitazionali.

Esistono analogie formali tra elettromagnetismo e gravitazione che hanno permesso la derivazione delle equazioni del campo gravitazionale. Risulta abbastanza naturale quindi che queste equazioni predicano l'esistenza di onde gravitazionali simili a quelle elettromagnetiche. La verifica sperimentale ha dovuto attendere fino al 1978, anno in cui il pulsar binario PS 1916+13 ha mostrato caratteristiche di emissione in accordo con la teoria.

È molto difficile ottenere soluzioni esatte delle equazioni del campo, data la loro non-linearità. Per questa ragione ci si limiterà ad una discussione delle equazioni linearizzate. La non-linearità trae origine dalla circostanza che il campo gravitazionale, possedendo energia ed impulso, è in realtà sorgente di se stesso. L'approssimazione lineare vale dunque solamente per onde cosí deboli da produrre un campo trascurabile.

Si sono già dedotte nel § 7.2 le equazioni statiche linearizzate di Einstein

$$\Box h_{\mu\nu} = -\frac{16G\pi}{c^4} S_{\mu\nu}, \tag{7.28}$$

dove

$$S_{\mu\nu} = T_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} T^{\lambda}_{\lambda}, \tag{7.29}$$

mentre le condizioni di armonicità sulle coordinate implicano

$$\frac{\partial}{\partial x^{\mu}_{\nu}} h^{\mu}_{\nu} = \frac{1}{2}\frac{\partial}{\partial x^{\nu}} h^{\mu}_{\mu}. \tag{7.30}$$

La (7.28) descrive la propagazione di onde nel vuoto con velocità c, qualora si ponga $S_{\mu\nu} = 0$. È molto istruttivo esaminarne le soluzioni mediante onde piane; l'interesse è per una soluzione del tipo

$$h_{\mu\nu} = e_{\mu\nu} \exp(ik_\lambda x^\lambda) + e^*_{\mu\nu} \exp(-ik_\lambda x^\lambda). \tag{7.31}$$

Dalla $\Box h_{\mu\nu} = 0$ segue

$$k_\lambda k^\lambda = 0 \tag{7.32.1}$$

e dalla (7.30)

$$k_\mu e^\mu{}_\nu = \frac{1}{2} k_\nu e^\mu{}_\mu. \tag{7.32.2}$$

Senza perdita di generalità si può supporre che l'onda sia diretta lungo l'asse x, ossia: $k^0 = k^1 = k$, $k^2 = k^3 = 0$. Dalla (7.32.2) segue: $e_{02} + e_{12} = 0$, $e_{03} + e_{13} = 0$, $e_{10} + e_{11} = -e_{00} - e_{01} = (e_{11} + e_{22} + e_{33} - e_{00})/2$.

La condizione di armonicità non fissa completamente le coordinate. È possibile infatti eseguire ulteriori trasformazioni del tipo $x^\alpha \to x^\alpha + \xi^\alpha$, dove

$$\xi^\alpha = i\varepsilon^\alpha \exp(ik_\lambda x^\lambda) - i\varepsilon^{*\alpha} \exp(-ik_\lambda x^\lambda), \tag{7.33}$$

in modo che le (7.28) e (7.30) siano ancora soddisfatte. Poiché la (7.33) implica la sostituzione $e_{\mu\nu} \to e_{\mu\nu} + k_\mu \varepsilon_\nu + k_\nu e_\mu$, è possibile scegliere ε_μ in modo da annullare e_{02}, e_{03}, e_{00}, e_{11}. Ne segue $e_{12} = e_{13} = 0$ ed inoltre: $e_{10} = (e_{22} + e_{33})/2 = -e_{10} = 0$. Sopravvivono solamente e_{23}, e_{22}, e_{33}, legate dalla $e_{22} + e_{33} = 0$. Dunque la generica onda piana contiene solamente due stati di polarizzazione, analogamente alle onde elettromagnetiche.

A parte questa analogia, uno studio piú dettagliato del comportamento degli stati di polarizzazione nei due casi, rispetto a rotazioni intorno all'asse x, metterebbe in evidenza che, mentre i fotoni hanno spin 1, ai gravitoni (cioè i quanti del campo gravitazionale) compete spin 2. In generale, una particella di spin 1 (2) ha 3 (5) stati di polarizzazione; nel caso dei fotoni la riduzione del numero di stati di polarizzazione segue dalla nullità della loro massa, mentre, nel caso dei gravitoni, la contrazione relativistica impedisce l'esistenza di stati il cui spin non è allineato con la direzione del moto.

Un gravitone porta con sé energia, che però non appare in $T_{\mu\nu}$, bensí quale termine non lineare nelle $R_{\mu\nu}$ stesse. In generale non è possibile separare in modo covariante la densità di energia e d'impulso di un gravitone dagli altri contributi del campo in $R_{\mu\nu}$. Tuttavia si può euristicamente associare all'onda (7.31) il tensore effettivo

$$\langle T_{\mu\nu} \rangle = c^2 \frac{k_\mu k_\nu}{8\pi G} (|e_{22}|^2 + |e_{23}|^2). \tag{7.34}$$

Questa è una formula molto utile per valutare la potenza emessa da sorgenti di onde gravitazionali. Cosí come una carica accelerata è necessaria per emettere onde elettromagnetiche, sono necessarie masse accelerate per emettere onde gravitazionali.

Il calcolo esatto di questa emissione si può effettuare risolvendo la (7.28) mediante potenziali ritardati ed usando la (7.34) per valutare il flusso energetico della sorgente.

Una stima approssimata si può ottenere dalla formula

$$P_e = \frac{(P_{in})^2}{P_0} \tag{7.35}$$

dove $P_0 = c^5/G = 3{,}69 \cdot 10^{59}$ erg/sec è una potenza «universale», e

$$P_{in} = \frac{MR^2}{T^3}, \tag{7.36}$$

ove si supponga che nel sistema emittente venga spostata una massa M nel tratto R, nel tempo T. Dato il valore enorme di P_0, la potenza P_e è sempre piccolissima, e questo spiega le enormi difficoltà finora incontrate nei tentativi di mettere in evidenza le onde gravitazionali.

Valga l'esempio di un corpo rotante, per il quale si ha

$$P_e = \frac{32 G \Omega^6 I^2 e^2}{5c^3} \tag{7.37}$$

dove Ω è la frequenza di rotazione, $I = I_{11} + I_{22}$, $e = (I_{11} - I_{22})/I$, e I_{11}, I_{22} sono i momenti di inerzia lungo gli assi ortogonali a quello di rotazione. Se il corpo è simmetrico rispetto a questo asse, non c'è emissione. Piú generalmente, è necessario un quadrupolo variabile per avere emissione di onde gravitazionali: non esistono né onde di monopolo, né di dipolo.

La (7.37) può applicarsi al caso di un pianeta di massa m in orbita alla distanza r. Si ha: $I_{22} = 0$, $e = 1$, $I_{11} = mr^2$, per cui

$$P_e = \frac{32 G \Omega^6 m^2 r^4}{5c}.$$

Per il pianeta Giove: $m = 1{,}9 \cdot 10^{30}$ g, $r = 7{,}78 \cdot 10^{13}$ cm, $\Omega = 1{,}68 \cdot 10^{-8}$ sec^{-1}, e quindi $P_e = 3{,}5$ kW, che è una potenza assolutamente trascurabile: occorrerebbero 10^{23} anni per registrare la caduta di Giove nel Sole. Nel caso del pulsar binario PS 1916+13, la separazione delle componenti è dell'ordine di 10^{11} cm, $m \sim 10^{33}$ g, ed il periodo è di alcune ore; la potenza emessa risulta dell'ordine di 10^{31} erg/sec e il sistema decade tipicamente in un miliardo di anni. L'irraggiamento cresce velocemente al ridursi della distanza tra le due componenti; alla separazione di 1000 km, si avrebbe $P_e \sim 10^{46}$ erg/sec e il collasso avverrebbe in meno di tre ore.

Come confronto si consideri una sbarra di acciaio del diametro di 1 metro e di lunghezza 20 metri, della massa di 490 tonnellate, che ruoti attorno ad un asse trasversale con una frequenza massima compatibile con la resistenza del materiale: $\Omega = 28$ rad/sec. La potenza radiata vale circa $2{,}2 \cdot 10^{-22}$ erg/sec, del tutto trascurabile rispetto alle perdite per attrito. La creazione di sorgenti di onde gravitazionali in laboratorio appare ancora poco probabile. Molto piú promettente è l'idea di Weber [1961] di rivelare onde gravitazionali provenienti da sorgenti astrofisiche (collassi da supernova) mediante le vibrazioni da esse indotte in un grande cilindro di alluminio. Un'altra idea di Braginsky [1965] utilizza delle cavità superconduttrici in cui la radiazione incidente dovrebbe indurre

delle transizioni nei modi di oscillazione. Un aumento nella sensibilità di questi detettori di pochi ordini di grandezza potrebbe aprire improvvisamente l'astronomia gravitazionale come nuovo campo d'indagine.

7.5. Le metriche isotrope.

Il calcolo di importanti effetti relativistici può effettuarsi solamente quando sia nota con precisione la metrica nell'intorno di un generico corpo a simmetria sferica. Einstein fu il primo ad ottenere nel 1915 delle soluzioni approssimate sotto forma di serie. La soluzione esatta fu data da Schwarzschild [1916]. Queste metriche possono venire scritte nel modo piú semplice introducendo coordinate quasi-minkowskiane x^1, x^2, x^3, $x^0 = t$, legate alle solite coordinate polari:

$$x^1 = r \cos\varphi \sin\vartheta \qquad x^2 = r \sin\varphi \sin\vartheta \qquad x^3 = r \cos\vartheta.$$

Allora si dirà isotropa una metrica del tipo

$$du^2 = B(r)\,dt^2 - A(r)\,dr^2 - r^2(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2). \tag{7.38}$$

In tal caso, le uniche componenti non nulle dei simboli di Christoffel sono

$$\begin{aligned} &\Gamma^r_{rr} = \frac{A'}{2A} && \Gamma^r_{\vartheta\vartheta} = -\frac{r}{A} && \Gamma^r_{\varphi\varphi} = -\frac{r \sin^2\vartheta}{A} \\ &\Gamma^r_{tt} = \frac{B'}{2A} && \Gamma^\vartheta_{r\vartheta} = \Gamma^\varphi_{r\varphi} = \frac{1}{r} && \Gamma^\vartheta_{\varphi\varphi} = -\sin\vartheta \cos\vartheta \\ &\Gamma^\varphi_{\varphi\vartheta} = \operatorname{ctg}\vartheta && \Gamma^t_{tr} = \frac{B'}{2B} \end{aligned} \tag{7.39}$$

avendo usato gli indici t, r, ϑ, φ. Pertanto

$$\begin{aligned} R_{rr} &= \frac{B''}{2B} + \frac{1}{4}\frac{B'}{B}\left(\frac{A'}{A} + \frac{B'}{B}\right) - \frac{1}{r}\frac{A'}{A} \\ R_{\vartheta\vartheta} &= -1 + \frac{r}{2A}\left(\frac{B'}{B} - \frac{A'}{A}\right) + \frac{1}{A} \\ R_{\varphi\varphi} &= \sin^2\vartheta\, R_{\vartheta\vartheta} \\ R_{tt} &= -\frac{B'}{2A} + \frac{1}{4}\frac{B'}{A}\left(\frac{A'}{A} + \frac{B'}{B}\right) - \frac{1}{r}\frac{B'}{A} \\ R_{\mu\nu} &= 0 \qquad \text{se } \mu \neq \nu. \end{aligned} \tag{7.40}$$

In tutte queste formule si è sottintesa la dipendenza esplicita di A e B da r; un apice denota differenziazione rispetto ad r. Dalla (7.25) si vede che nello spazio vuoto si ha

$$R_{\mu\nu} - g_{\mu\nu} R/2 = 0, \tag{7.41}$$

che, contratta con $g^{\mu\nu}$, dà $R = 0$, $R_{\mu\nu} = 0$. Utilizzando le (7.40) si trova

$$B = 1 - \frac{2MG}{c^2 r} \qquad AB = 1 \tag{7.42}$$

dove M è una costante d'integrazione. Tuttavia, dato che a grandi distanze ci si aspetta che la soluzione descriva il solito potenziale newtoniano, confrontando B con la (7.7) risulta che la costante M è da identificarsi con la massa del corpo. Perciò la metrica di Schwarzschild è data da

$$(7.43)\qquad du^2=\left(1-\frac{2MG}{c^2r}\right)dt^2-\left(1-\frac{2MG}{c^2r}\right)^{-1}dr^2-r^2(d\vartheta^2+\sin^2\vartheta\, d\varphi^2).$$

Esistono altre forme di questa metrica ad essa equivalenti [cfr. Weinberg 1972; Misner, Thorne e Wheeler 1973]. Si noterà che la (7.43) diventa degenere per $r=2MG/c^2$, noto come raggio di Schwarzschild. Non si tratta tuttavia di un punto singolare della soluzione, bensí di una singolarità fisicamente inessenziale, derivante dall'uso di una carta di coordinate inadeguata.

Piú recentemente, Finkelstein [1958] e Kruskal [1960] sono giunti a formulare delle nuove coordinate che chiariscono completamente la struttura geometrica della soluzione (7.43). Usando la lunghezza $\mathfrak{M}\equiv MG/c^2$ come misura della massa del corpo, si introducano le variabili

$$(7.44.1)\qquad z=\left|1-\frac{r}{2\mathfrak{M}}\right|^{1/2}e^{t/4\mathfrak{M}}\sinh\left(\frac{t}{4\mathfrak{M}}\right)$$

$$(7.44.2)\qquad v=\left|1-\frac{r}{2\mathfrak{M}}\right|^{1/2}e^{t/4\mathfrak{M}}\cosh\left(\frac{t}{4\mathfrak{M}}\right)$$

per cui

$$(7.45)\qquad v^2-z^2=\left(1-\frac{r}{2\mathfrak{M}}\right)e^{t/2\mathfrak{M}}.$$

Si può considerare $r=r(z, v)$ come una funzione di z, v, ottenuta risolvendo l'equazione implicita (7.45). Sostituendo nella (7.43) si ottiene l'espressione

$$(7.46)\qquad du^2=\frac{32\mathfrak{M}^3}{r}e^{-r/2\mathfrak{M}}(dv^2-dz^2)-r^2(d\vartheta^2+\sin^2\vartheta\, d\varphi^2),$$

che non presenta alcuna singolarità in $r=2\mathfrak{M}$.

La corrispondenza tra t, r, e z, v, dà il quadrante Q_1, $z>0$ e $|v|<z$, come immagine di $r>2\mathfrak{M}$ e $-\infty<t<+\infty$. D'altra parte, la metrica (7.46) ha senso anche in tutta la regione Q: $v^2-z^2<1$, nel piano z, v, che contiene Q_1. Tale regione estesa contiene anche il quadrante Q_2, $z<0$ e $|v|<-z$, simmetrico rispetto al primo. La frontiera $v^2-z^2=1$ costituisce una vera singolarità per la soluzione in $r=0$. In Q, oltre Q_1, Q_2, vi sono anche le regioni R_1: $|z|<v<\sqrt{z^2+1}$, e R_2: $-\sqrt{z^2+1}<v<-|z|$. Poiché il cono luce $du=0$ – restringendosi a ϑ, φ costanti – si presenta nel piano z, v come $z\pm v=\text{cost}$, si vede che le traiettorie dei corpi materiali non possono deviare piú di 45° dal semiasse positivo di v. Pertanto Q_1 e Q_2 non possono comunicare fra loro, mentre R_1 può ricevere segnali da Q_1, Q_2, ma non può inviarne ad essi, mentre R_2 può inviare segnali a Q_1, Q_2, ma non riceverne.

Pertanto la soluzione di Schwarzschild rappresenta il campo esterno alla massa, e può essere continuata analiticamente, per mezzo della (7.46), anche

nella regione $r<2\mathfrak{M}$, là dove essa è vuota. Poiché da questa regione non possono essere inviati segnali all'esterno, essa viene chiamata «buco nero».

Fisicamente, se una data massa viene compressa in un raggio sempre piú piccolo, si possono realizzare estensioni di spazio vuoto che si avvicinano sempre piú a $r=2\mathfrak{M}$. Successivamente, la struttura diventa instabile e la forza gravitazionale vince tutte le altre; il sistema procede verso il collasso gravitazionale e si forma un buco nero. La zona esterna alla massa in collasso è ancora rappresentata da un frammento del quadrante Q. Un osservatore A sulla superficie della stella finisce per raggiungere la sfera $r=2\mathfrak{M}$ ed entrare in R_1. Per un osservatore esterno B, l'orizzonte $r=2\mathfrak{M}$ viene raggiunto da A in un tempo infinito e l'entrata in R_1 non viene mai osservata. A invece impiega un tempo finito dell'ordine di $2\mathfrak{M}/c$ per arrivare su $r=2\mathfrak{M}$; giunto in R_1, A deve procedere di necessità fino alla singolarità $r=0$, $v=\sqrt{z^2+1}$, dove le equazioni del campo cessano di avere significato. Un frammento di regione R_1 è dunque effettivamente raggiungibile attraverso il collasso gravitazionale; varcato tuttavia l'orizzonte $r=2\mathfrak{M}$, non sarà piú possibile comunicare ad un osservatore esterno B i risultati dell'osservazione.

7.6. Moto di un corpo nel campo isotropo.

Verrà ora discusso il moto geodetico di un corpo di prova nel campo esterno a una massa sferica, e si applicheranno i risultati al caso del Sole. Le equazioni della caduta libera sono date dalle (7.2); introducendo le componenti di Christoffel (7.39), si può scrivere esplicitamente:

$$(7.47.1)\qquad \frac{d^2r}{du^2}+\frac{A'}{2A}\left(\frac{dr}{du}\right)^2-\frac{r}{A}\left(\frac{d\vartheta}{du}\right)^2-r\frac{\sin^2\vartheta}{A}\left(\frac{d\varphi}{du}\right)^2+\frac{B'}{2A}\left(\frac{dt}{du}\right)^2=0$$

$$(7.47.2)\qquad \frac{d^2\vartheta}{du^2}+\frac{2}{r}\frac{d\vartheta}{du}\frac{dr}{du}-\sin\vartheta\cos\vartheta\left(\frac{d\varphi}{du}\right)^2=0$$

$$(7.47.3)\qquad \frac{d^2\varphi}{du^2}+\frac{2}{r}\frac{d\varphi}{du}\frac{dr}{du}+2\operatorname{ctg}\vartheta\frac{d\varphi}{du}\frac{d\vartheta}{du}=0$$

$$(7.47.4)\qquad \frac{d^2t}{du^2}+\frac{B'}{B}\frac{dt}{du}\frac{dr}{du}=0.$$

Queste equazioni si discutono cercando le costanti del moto, e stabilendo in quale approssimazione esse si riducono al moto kepleriano nel potenziale di Newton. Converrà notare subito che, causa la isotropia della metrica, non si ha perdita di generalità supponendo $\vartheta=\pi/2$; la (7.47.3) risulta allora identicamente soddisfatta. La (7.47.1) equivale a $\frac{d}{du}\left(r^2\frac{d\varphi}{du}\right)=0$, ossia

$$(7.48)\qquad r^2\frac{d\varphi}{du}=\frac{\mathcal{J}}{c}=\text{cost}.$$

Si ottiene in questo modo l'analogo della legge delle aree o conservazione del momento angolare. Inoltre, dalla (7.47.4),

$$\frac{d}{du}\left(B\frac{dt}{du}\right)=0,$$

da cui

$$\frac{dt}{du}=\frac{E}{B} \qquad E=\text{cost}, \tag{7.49}$$

che, come si vedrà subito, esprime la conservazione dell'energia. Introducendo (7.48) e (7.49) nella (7.47.1), si ottiene

$$\frac{d}{du}\left\{A\left(\frac{dr}{du}\right)^2+\frac{\mathcal{J}^2}{c^2r^2}-\frac{E^2}{B}\right\}=0,$$

per cui si ha la costante del moto

$$A\left(\frac{dr}{du}\right)^2+\frac{\mathcal{J}^2}{c^2r^2}-\frac{E^2}{B}=-K. \tag{7.50}$$

Si supponga ora $K\neq 0$; poiché le (7.2) sono invarianti se si moltiplica u per un fattore costante non nullo (trasformazioni di scala), si può anche supporre $K=1$.

Per ottenere la corretta interpretazione fisica delle costanti E, $\mathcal{J}$, occorre rifarsi al caso kepleriano, in cui le velocità sono basse e $r\gg 2\mathfrak{M}$. Conviene reintrodurre la variabile temporale T: $x^0=t=cT$. Per un moto nel piano x,y, il momento angolare è dato da $r^2(d\varphi/dt)$ $(m=1)$, che coincide con $\mathcal{J}/c$ nel limite considerato. Se il corpo è in un'orbita aperta, all'infinito si ha, per la (7.42), $B\sim 1$, e $dt\sim du$ (moto lento), per cui E deve essere molto prossimo a 1. Nel caso kepleriano la conservazione dell'energia è data da:

$$\frac{1}{2}\left\{\left(\frac{dr}{dT}\right)^2+\frac{\mathcal{J}^2}{r^2}-\frac{GM}{r^2}\right\}=\varepsilon. \tag{7.51}$$

La (7.51) coincide con la (7.50) se si tiene conto che $u=cT$ e che, per la (7.42), $A\simeq 1$; inoltre bisogna porre $E\simeq 1+\varepsilon/c^2$, dove ε è l'energia non relativistica del moto. In generale, per orbite aperte, si ha $E=1/\sqrt{1-\beta_\infty^2}$, dove β_∞ va calcolato all'infinito. E corrisponde dunque all'energia del sistema.

Il calcolo dell'orbita segue gli stessi procedimenti del caso kepleriano, e si imposta eliminando il tempo ed introducendo φ come variabile indipendente, usando la (7.48). Si ottiene

$$\frac{1}{r^4}\left(\frac{dr}{d\varphi}\right)^2+\frac{1}{r^2}-\frac{2MG}{c^2r^3}+\frac{1-E^2}{\mathcal{J}^2}c^2-\frac{2MG}{\mathcal{J}^2r}=0,$$

che, introducendo le variabili adimensionate $I=\mathcal{J}c/MG$, $z=MG/rc^2$, si riduce a

$$\left(\frac{dz}{d\varphi}\right)^2=\frac{E^2-(1-2z)(1+I^2z^2)}{I^2}. \tag{7.52}$$

Per le orbite circolari vale immutata la legge di Keplero: $\omega^2r_0^3=MG$, dove $\omega=d\varphi/dT$.

È impossibile integrare esattamente queste equazioni del moto nel caso generale senza ricorrere alle funzioni ellittiche. Per orbite quasi circolari si ponga invece: $z(\varphi)=z_0+\eta(\varphi)$, con $3z_0^2-z_0+1/I^2=0$; derivando la (7.52), si ha allora

$$\frac{d^2\eta}{d\varphi^2}+(1-6z_0)\eta=3\eta^2$$

e, trascurando il termine η^2, si ottiene la soluzione approssimata

$$\eta(\varphi)=\eta_0\cos(\sqrt{1-6z_0}\,\varphi+\varphi_0),$$

dove η_0, φ_0 sono costanti d'integrazione. Due massimi successivi in $\eta(\varphi)$ sono intervallati in φ da $\Delta=2\pi/\sqrt{1-6z_0}$; le orbite quasi circolari non si chiudono in se stesse come per le ellissi di Keplero $(\Delta=2\pi)$. Per piccoli z, si ha: $\Delta\varphi\equiv\equiv\Delta-2\pi\cong 6\pi GM/c^2r_0$, che costituisce un'ottima approssimazione per i pianeti dotati notoriamente di una piccola precessione. Per Mercurio, $\Delta\varphi$ vale 0,1038″ per rivoluzione, circa 43,03″ per secolo. La teoria newtoniana darebbe come precessione, causata dalla perturbazione dei pianeti, $\Delta\varphi_N=5557{,}62''$/ secolo; si osserva invece: $\Delta\varphi_{\text{oss}}=5600{,}73''$/secolo, e la differenza $\Delta\varphi_N-\Delta\varphi_{\text{oss}}$ è in ottimo accordo con la predizione $\Delta\varphi$ einsteiniana.

A questo risultato, di notevole importanza non solamente storica, non sono mancate critiche. Uno schiacciamento equatoriale del Sole condurrebbe a deviazioni quadrupolari dalla legge di Newton, che potrebbero influenzare il calcolo classico di $\Delta\varphi_N$. Questo lascerebbe spazio per la teoria di Jordan [1959] e Brans-Dicke [1961], una generalizzazione della relatività generale. Altri dati osservativi conducono a ritenere sempre meno probabile questa teoria e trascurabile lo schiacciamento solare. Anche gli altri pianeti (in particolare la Terra, Venere ed Icaro) mostrano precessioni in accordo con la teoria, ma soggette ad errori osservativi piú rilevanti. Spettacolare invece il già citato pulsar binario PS 1916+13, che mostra una forte precessione causata dall'orbita molto eccentrica e stretta (il semiasse maggiore è dell'ordine di un milione di chilometri) ed in perfetto accordo con le previsioni relativistiche.

Rimane da considerare il caso della deflessione della luce nel campo gravitazionale di una stella di massa M. Si ricordi che, per $m=1$, $c^2E=c^2/\sqrt{1-\beta_\infty^2}$; inoltre, detto b il parametro d'impatto, cioè la distanza tra il centro della stella e la retta cui tende il corpo, si ha $\mathcal{J}=bp_\infty$, dove $p_\infty=\beta_\infty c/\sqrt{1-\beta_\infty^2}$ è il momento lineare all'infinito. Pertanto, nel limite $v\to c$, si ha $\mathcal{J},I,E\to\infty$, mentre $I/E\to\to b/\mathfrak{M}$. La (7.52) diventa

$$\left(\frac{dz}{d\varphi}\right)^2=\left(\frac{\mathfrak{M}}{b}\right)^2-z^2(1-2z) \tag{7.53}$$

e, differenziando,

$$\frac{d^2z}{d\varphi^2}+z=3z^2. \tag{7.54}$$

In assenza del termine in z^2, si avrebbe l'orbita rettilinea parallela all'asse x

$$z_0(\varphi)=\frac{\mathfrak{M}}{b}\sin\varphi;$$

per piccole deviazioni da essa, si ponga $z(\varphi)=z_0(\varphi)+z_1(\varphi)$, con

$$\frac{d^2z_1}{d\varphi^2}+z_1=3z_0^2=3\left(\frac{\mathfrak{M}}{b}\right)^2\sin^2\varphi,$$

e condizioni iniziali $z(0)=0$, $dz/d\varphi|_0=\mathfrak{M}/b$. Allora si ottiene la soluzione approssimata

$$z(\varphi)=\frac{\mathfrak{M}}{b}\sin\varphi+4\left(\frac{\mathfrak{M}}{b}\right)^2\left(\sin\frac{\varphi}{2}\right)^4,$$

che possiede, oltre allo zero di «entrata» in $\varphi=0$, anche lo zero di «uscita» per $\varphi=\pi+4\mathfrak{M}/b+O(\mathfrak{M}/b)^2$. In questa approssimazione, la deflessione è dunque $4MG/bc^2$.

Per il Sole, questa quantità vale circa $1{,}75''$, predizione verificata per la prima volta durante un'eclissi di Sole a Sobral nel 1919; altre verifiche furono fatte durante successive eclissi fino al 1952 in Sudan. La conferma dell'effetto rese Einstein famosissimo presso il grande pubblico. Il procedimento usato richiede il confronto tra due fotografie dello stesso campo stellare, una ripresa vicino al bordo solare durante l'eclissi, l'altra a distanza di sei mesi in assenza del Sole. Tutte le misure ottiche effettuate finora concordano con errori non superiori al 20 per cento circa.

Migliori risultati si ottengono misurando la deflessione di radioonde emesse da sorgenti quasi-stellari mediante radiotelescopi. Due di queste sorgenti, 3C279 e 3C273, vengono periodicamente occultate dal Sole e si prestano a misure che risultano affette da errori molto minori (1 per cento). Altre verifiche implicano la misura di echi radar da parte di pianeti e satelliti artificiali in orbita circumsolare.

7.7. Generalità sul collasso gravitazionale.

La soluzione di Schwarzschild è stata generalizzata al caso di un corpo rotante dotato di carica elettrica, secondo la soluzione di Kerr-Newman. Di questa, date le notevoli complicazioni tecniche, non si riporterà che la formula finale:

$$ds^2=-\frac{\Delta}{\rho^2}(dt-a\sin^2\vartheta\,d\varphi)^2+ \tag{7.55}$$
$$+\frac{\sin^2\vartheta}{\rho^2}[(r^2+a^2)\,d\varphi-a\,dt]^2+\frac{\rho^2}{\Delta}dr^2+\rho^2d\vartheta^2,$$

dove $\rho^2=r^2+a^2\cos^2\vartheta$, e $\Delta=r^2-2\mathfrak{M}r+a^2+Q^2$. Il momento angolare del corpo è dato da $\mathcal{J}=Mac$, e la carica elettrica è $Q=e\sqrt{8\pi G/c^4}$. Esiste dunque an-

che un campo elettromagnetico con un dipolo magnetico indotto dalla rotazione; esso ha la forma:

$$F=\frac{1}{2}F_{\mu\nu}dx^\mu\wedge dx^\nu=Q\rho^{-4}\{(r^2-a^2\cos^2\vartheta)\,dr\wedge \tag{7.56}$$
$$\wedge(dt-a\sin^2\vartheta\,d\varphi)+2ar\cos\vartheta\sin\vartheta\,d\vartheta\wedge[(r^2+a^2)\,d\varphi-a\,dt]\},$$

dove, ad esempio, $dx^\mu\wedge dx^\nu=-dx^\nu\wedge dx^\mu$. Questo campo contribuisce al tensore energia-impulso $T_{\mu\nu,EM}$.

La Kerr-Newman è molto importante per lo studio di collassi gravitazionali non isotropi. Su ulteriori generalizzazioni di questa soluzione non esistono teoremi rigorosi se non quello di Birkhoff e la congettura del *no hair*. Il primo è l'analogo del risultato di Newton, secondo cui ogni potenziale gravitazionale a simmetria sferica è necessariamente un multiplo di quello newtoniano:

TEOREMA (Birkhoff). *Se la geometria di una regione dello spazio-tempo ha simmetria sferica ed è soluzione delle equazioni di Einstein nel vuoto, allora la regione è necessariamente un frammento di quella di Schwarzschild.*

Per la prova si rimanda a Birkhoff [1923]. Esso implica l'assenza di onde gravitazionali a simmetria sferica e quindi l'assenza di onde nel collasso gravitazionale isotropo. Si dà molto credito anche alla congettura *no hair*, secondo cui lo stadio finale del collasso gravitazionale di un oggetto è sempre rappresentato da una soluzione di Kerr-Newman. Va notato che esistono altre soluzioni statiche delle equazioni del campo. Tutte queste soluzioni mostrano singolarità cosiddette «nude», tali cioè da essere in principio osservabili dall'esterno. Nella Kerr-Newman, invece, queste singolarità (come si è visto nel caso particolare della Schwarzschild) sono poste al di là dell'orizzonte degli eventi e quindi non sono osservabili.

Lo stadio finale del collasso dovrebbe essere caratterizzato unicamente da M, Q, $\mathcal{J}$, e quindi da parametri rilevabili mediante osservazioni effettuate con campi a lungo raggio d'azione, quali il campo gravitazionale e quello elettromagnetico. La formazione del buco nero cancella tutta l'informazione riguardante gli altri campi o le eventuali irregolarità ed anisotropie iniziali del corpo da cui ha tratto origine. Esso appare, o appariva, come la morte ultima della materia.

È possibile estrarre, in teoria, energia dalla Kerr-Newman secondo un esperimento concettuale di Penrose [1969] a spese dell'energia rotazionale del buco. Quando questa sia esaurita, si giunge alla soluzione isotropa da cui non è possibile estrarre energia con mezzi basati sulla meccanica classica.

Negli ultimi anni sono state studiate molte proprietà dei buchi neri, che stabiliscono interessanti legami con la meccanica dei quanti e con il secondo principio della termodinamica, che è stato generalizzato. Di queste ricerche si daranno solo alcune considerazioni euristiche.

Poiché la creazione di un buco nero distrugge informazione in modo irreversibile, si può assegnare al buco nero un certo ammontare di entropia proporzionale al volume dell'oggetto. L'entropia di un insieme di buchi neri può solamente crescere a seguito delle loro interazioni. Se due o piú di essi collidono, possono

coalescere formandone uno con un volume totale eccedente quello dei frammenti. Hawking [1975] ha tuttavia dimostrato che un buco nero può effettuare una transizione quantistica emettendo fotoni (ma anche altre particelle) e diminuire di massa fino a scomparire. L'emissione di fotoni avviene secondo lo spettro del corpo nero di Planck, con una temperatura che è inversamente proporzionale alla massa del buco e con una intensità che va come $\hbar c^6/G^2M^2$. Procedendo l'evaporazione, aumentano dunque la temperatura del buco e la potenza di emissione, fino al punto in cui questo scompare in una esplosione. La vita media di un buco nero di 10^{15} grammi è valutata circa 10^{10} anni, l'età cioè dell'Universo. Oggetti di questo tipo formatisi durante il *big bang* esploderebbero adesso, e potrebbero venire osservati. I buchi neri formati dal collasso di masse stellari avrebbero invece una vita media troppo lunga (10^{64} anni) e comunque riceverebbero piú energia dalla radiazione fossile, di quanta ne perderebbero evaporando. L'entropia del buco nero si ritrova sotto forma di entropia del gas di fotoni risultante dall'evaporazione; si generalizza in questo modo il secondo principio della termodinamica.

Quali sono i dati osservativi in nostro possesso sull'esistenza effettiva dei buchi neri? La sorgente Cygnus X 1, a raggi X, sembra avere molte caratteristiche spiegabili solamente con un modello di accrezione, in cui un buco nero è in orbita molto stretta attorno a una gigante blu. L'atmosfera di questa gigante blu precipita nel campo del buco nero, viene soggetta a violenta compressione e riscaldamento, dando luogo ad emissione di raggi X, effettivamente osservati con le caratteristiche predette dalla teoria. Esistono altre spiegazioni che non implicano l'esistenza del buco nero; tuttavia esse appaiono piú artificiali di quella proposta, e saranno probabilmente scartate con l'affinarsi sia della teoria, sia dell'osservazione.

La distribuzione delle stelle nelle galassie giganti ellittiche quali M87, in Virgo, mostra delle strutture anomale a cuspide da cui proviene una emissione di raggi X, e da cui pare uscire un getto di materia ad alta velocità. È molto difficile immaginare delle sorgenti termonucleari capaci di provocare fenomeni del genere; infatti queste reazioni convertono al piú lo 0,8 per cento della massa totale in energia, mentre è possibile che il buco nero si dimostri un convertitore molto piú efficiente. Si pensa dunque che nel centro di questa galassia esista un buco nero con una massa di circa 10^9 masse solari. Oggetti del genere esisterebbero, ma di dimensioni minori, nel centro degli ammassi globulari.

Non s'insisterà oltre su questo soggetto, di alto interesse, sia per ragioni di spazio sia perché lo sviluppo delle tecniche osservative è tale da portare a una radicale revisione delle nostre conoscenze nel giro di pochi anni, rendendo questa rassegna rapidamente obsoleta.

Occorre comunque avvertire il lettore che la presente esposizione della relatività non pretende di essere esauriente, in quanto non tratta, ad esempio, argomenti molto importanti di astrofisica e cosmologia, strettamente basati sulla teoria della relatività generale [per essi si rimanda ancora una volta a Weinberg 1972; Misner, Thorne e Wheeler 1973; e ad altri articoli di questa stessa *Enciclopedia*]. [T. R.].

Birkhoff, G. D.
1923 *Relativity and Modern Physics*, Harvard University Press, Cambridge Mass.

Braginsky, V. B.
1965 *Gravitational radiation and the prospect of its experimental discovery*, in «Uspechi Fizičeskich Nauk», LXXXVI, pp. 433-46.

Brans, C., e Dicke, R. H.
1961 *Mach's principle and a relativistic theory of gravitation*, in «Physical Review», CXXIV, pp. 925-35.

Cartan, E.
1923-24 *Sur les variétés à connexion affine et la théorie de la Rélativité Généralisée*, in «Annales de l'Ecole Normale Supérieure», XL, pp. 325-412, e XLI, pp. 1-25.

Christoffel, E. B.
1869 *Über die Transformation der homogenen Differentialausdrücke zweiten Grades*, in «Journal für die reine und angewandte Mathematik», LXX, pp. 46-70.

Dirac, P.-A.-M.
1930 *The Principles of Quantum Mechanics*, Clarendon Press, Oxford 1958[4] (trad. it. Boringhieri, Torino 1959).

Dyson, F. W.; Eddington, A. S.; e Davidson, C.
1920 *A determination of the deflection of the light by the sun's gravitational field*, in «Philosophical Transactions of the Royal Society of London», serie A, CCXX, pp. 291-98.

Einstein, A.
1905*a* *Zur Elektrodynamik bewegter Körper*, in «Annalen der Physik», XVII, pp. 891-921.
1905*b* *Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?*, in «Annalen der Physik», XVIII, pp. 639-41.
1916 *Die Grundlage der allgemeine Relativitätstheorie*, in «Annalen der Physik», XLIX, pp. 769-822 (trad. it. in M. Pantaleo (a cura di), *Cinquant'anni di relatività*, Editrice Universitaria, Firenze 1955).
1922 *Vier Vorlesungen über Relativitätstheorie*, Vieweg, Braunschweig (trad. it. Boringhieri, Torino 1968[2]).

Eötvös, R. V.
1889 *Über die Anziehung der Erde auf verschiedene Substanzen*, in «Mathematische und naturwissenschaftliche Berichte aus Ungarn», VIII, pp. 65-68.

Finkelstein, D.
1958 *Past-future asymmetry of the gravitational field of a point particle*, in «Physical Review», CX, pp. 965-67.

Gauss, C. F.
1828 *Disquisitiones generales circa superficies curvas*, Dieterich, Göttingen.

Goldstein, H.
1959 *Classical Mechanics*, Addison-Wesley, Reading Mass. (trad. it. Zanichelli, Bologna 1971).

Hawking, S. W.
1975 *Particle creation by black holes*, in «Communications in Mathematical Physics», XLIII, pp. 199-220.

Holton, G.
1960 *On the origins of the special theory of relativity*, in «American Journal of Physics», XXVIII, pp. 627-43.

Hubble, E.
1929 *A relation between distance and radial velocity among extragalactic nebulae*, in «Proceedings of the National Academy of Sciences of the United States of America», XV, pp. 169-73.

Jordan, P.
1959 *Zum gegenwärtigen Stand der Diracschen Kosmologischen Hypothesen*, in «Zeitschrift für Physik», CLVII, pp. 112-21.

Kruskal, M. D.
1960 *Maximal Extension of Schwarzschild Metric*, in «Physical Review», CXIX, 5, pp. 1743-45.

Levi-Civita, T.
1917 *Nozione di parallelismo in una varietà qualunque e conseguente specificazione geometrica della curvatura riemanniana*, in «Rendiconti del Circolo Matematico di Palermo», XLII, pp. 173-204.

Lipschitz, R.
1869 *Untersuchungen in Betreff der ganzen homogenen Functionen von n Differentialen*, in «Journal für die reine und angewandte Mathematik», LXX, pp. 71-102.

Lorentz, H.
1904 *Electromagnetic phenomena in a system moving with any velocity smaller than that of light*, in «Atti dell'Academia Scientiarum di Amsterdam», VI, pp. 809-32; ora in H. Lorentz, A. Einstein, H. Minkowski e H. Weyl (a cura di), *The Principle of Relativity*, Dover, New York 1952.

Minkowski, H.
1908 *Die Grundgleichungen für die elektromagnetischen Vorgänge in bewegten Körpern*, in «Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-physikalische Klasse», I, pp. 53-111; ora in H. Lorentz, A. Einstein, H. Minkowski e H. Weyl (a cura di), *The Principle of Relativity*, Dover, New York 1952.

Misner, C. W.; Thorne, K. S.; e Wheeler, J. A.
1973 *Gravitation*, Freeman, San Francisco.

Pauli, W.
1921 «Relativitätstheorie», in *Encyclopädie der Mathematischen Wissenschaften*, vol. V, art. 19, Teubner, Leipzig, pp. 539-775 (trad. it. Boringhieri, Torino 1978²).

Penrose, R.
1969 *Gravitational Collapse: the Role of General Relativity*, in «Rivista del Nuovo Cimento», I, 252-76.

Poincaré, H.
1904 *The Principles of Mathematical Physics*, in H. J. Rogers (a cura di), *Proceedings of the Congress of Arts and Science*, vol. I, Houghton Mifflin, Boston 1905; ora in L. P. Williams (a cura di), *Relativity Theory: Its Origins and Impact on Modern Thought*, Wiley, New York 1968.

Rebka, G. A., e Pound, R. V.
1960 *Apparent Weight of Photons*, in «Physical Review Letters», IV, pp. 337-41.

Ricci, G., e Levi-Civita, T.
1901 *Méthodes de calcul différentiel absolu et leurs applications*, in «Mathematische Annalen», LIV, pp. 135-201.

Riemann, B.
[1854] *Über die Hypothesen, welche der Geometrie zu Grunde liegen...*, Dieterich, Göttingen 1867 (trad. it. in A. Einstein, *Relatività. Esposizione divulgativa*, Boringhieri, Torino 1967, pp. 204-20).

Schilpp, P. A.
1949 (a cura di) *Albert Einstein: Philosopher-Scientist*, The Library of Living Philosophers, Evanston Ill. (trad. it. Boringhieri, Torino 1958).

Schwarzschild, K.
1916 *Über das Gravitationsfeld eines Massenpunktes nach der Einsteinsche Theorie*, in «Sitzungsberichte der deutschen Akademie der Wissenschaften», CCCCXXIV, pp. 189-96.

Weber, J.
1961 *General Relativity and Gravitational Waves*, Wiley, New York.

Weinberg, S.
1972 *Gravitation and Cosmology*, Wiley, New York.

Whittaker, E.
1953 *A History of the Theories of Aether and Electricity*, Nelson, London.

Woolf, H.
1979 (a cura di) *Albert Einstein Centennial Celebration*, Addison-Wesley, Reading Mass.

Le idee fondamentali della relatività si sono innestate su una profonda revisione delle nozioni di spazio e di tempo (cfr. **spazio-tempo**) e di **moto** per arrivare infine a una sofisticata modellistica (cfr. **modello**, **teoria/modello**) dell'**universo**, particolarmente rilevante per le diverse **cosmologie** attuali. All'elaborazione einsteiniana del 1905 che, com'è noto, conteneva i punti di fondo della cosiddetta relatività speciale, che veniva sempre piú autorevolmente a rappresentare il quadro concettuale (cfr. **paradigma**) adatto allo studio del campo elettromagnetico (cfr. **forza/campo**), si è affiancata a partire dal 1916 l'elaborazione della relatività «generale« che ha invece prospettato sotto forma nuova la questione della **gravitazione**. Lo stesso Einstein, infine, ha elaborato una teoria del campo unificato, avanzando l'esigenza di un quadro concettuale unitario (cfr. **uno/molti**) per le interazioni (cfr. **interazione**) fisiche fondamentali. Nella **fisica** attuale la relatività speciale si ritiene ormai ampiamente corroborata; piú sottile è invece la questione delle «prove sperimentali» (cfr. **empiria/esperienza**, **esperimento** e, per altri aspetti, **equivalenza**) della relatività generale (una classica «prova» addotta è, notoriamente, l'anomalia del perielio di Mercurio: cfr. **pianeti**, **astronomia**), che si presenta non solo come uno sviluppo delle idee matematiche di Riemann (cfr. **matematiche**; e, per temi piú specifici, **geometria e topologia**, **sistemi di riferimento**), ma come un ripensamento radicale delle nozioni fisiche di **materia**, **luce**, ecc. Sotto quest'ultimo profilo la relatività è non solo un esempio di riuscita **rivoluzione** scientifica (cfr. **ricerca**, **scienza**) che ha visto infine prevalere le proprie modalità di **spiegazione** dei fenomeni, ma, forse, una concezione della **natura** (cfr. anche **mondo**) con cui le stesse filosofie e metafisiche hanno da confrontarsi (cfr. **concetto**, **idea**; **metafisica** e **filosofia/filosofie**).

Reversibilità/irreversibilità

Sono innumerevoli gli aspetti dell'esperienza quotidiana che contribuiscono a fornire solide basi empiriche all'idea che tutti i fenomeni naturali procedono in una ben definita direzione temporale, dal passato al presente e verso il futuro. È quindi profondamente radicata nella nostra coscienza una concezione del tempo la cui caratteristica fondamentale è la proprietà di flusso unidirezionale, con la naturale conseguenza di una profonda differenziazione tra il concetto di un passato irrimediabilmente trascorso, in cui tutto è assolutamente invariabile e di cui al piú si può aver memoria, e di un futuro ancora da realizzarsi, in cui tutto è completamente indeterminato e che si può in qualche modo influenzare, mediante l'intervento della nostra volontà e delle nostre azioni.

Tuttavia, nell'ambito di una piú particolare analisi scientifica del concetto di tempo, non è possibile assumere per scontata questa proprietà di assoluta intrinseca irreversibilità, per quanto forti possano essere gli argomenti in suo favore, sia di tipo fisiologico e psicologico sia di tipo storico-sociale o piú generalmente collegati a concezioni filosofiche. Infatti è ben noto che tutte le leggi fisiche elementari del moto, come quelle della meccanica classica, dell'elettrodinamica, della meccanica quantistica e della relatività generale, sono perfettamente reversibili rispetto al tempo, nel senso che ad ogni fenomeno elementare, che si svolge in una determinata direzione temporale rispettando le leggi del moto, corrisponde, in un senso ben definito, un altro fenomeno elementare che si svolge secondo la direzione temporale opposta ma ancora nel pieno rispetto delle leggi del moto.

Per dare un'immagine semplice ed immediata che serva ad illustrare questa contrapposizione si possono considerare le seguenti situazioni. Si supponga di effettuare una ripresa cinematografica di un qualsiasi evento di vita quotidiana, quale per esempio l'accensione di un fiammifero, lo sciogliersi al sole di un pupazzo di neve, il tuffo di un nuotatore dal trampolino in piscina... La successiva proiezione della pellicola riprodurrà sullo schermo l'evento, cosí come si è svolto, ma se s'immagina di proiettare la pellicola in senso inverso allora quanto appare sullo schermo sarà immediatamente riconosciuto, nella maggior parte dei casi, come innaturale e qualificato come impossibile a verificarsi nella realtà. Se invece la pellicola contiene la registrazione di fenomeni elementari che si svolgono secondo le leggi della meccanica, quali l'urto tra sfere perfettamente elastiche oppure la riproduzione del moto dei pianeti nel sistema solare, allora non sarà possibile rilevare differenze di «naturalezza» tra gli eventi riprodotti sullo schermo in corrispondenza dei due versi di proiezione della pellicola. Entrambi gli eventi infatti, nonostante la loro diversità, sembreranno svolgersi nel rispetto delle leggi del moto.

In generale, quindi, ogni processo fisico elementare può aver luogo in senso diretto, svolgendosi dal passato verso il futuro, oppure in senso temporale invertito, dal futuro verso il passato. Di conseguenza, per i processi fisici elemen-

tari non esiste una direzione privilegiata per il flusso temporale, poiché entrambe le direzioni, sia quella naturale, data dalla nostra sensazione empirica, sia quella inversa, sono perfettamente equivalenti nella descrizione dei fenomeni. Ci si trova allora di fronte a una drastica dicotomia tra il concetto di irreversibilità del senso del tempo, che per le sue solide basi fisiologiche, psicologiche, storiche e filosofiche, è profondamente connaturato alla nostra cultura e alla nostra civiltà, e il concetto opposto di reversibilità, che è alla base di tutte le leggi fisiche elementari del moto la cui validità sperimentale è solidamente stabilita nell'ambito delle varie classi di fenomeni trattati.

D'altra parte anche dal punto di vista fisico esistono casi importanti dove è possibile fare distinzione tra i due sensi di flusso del tempo. Infatti per tutti i processi fisici macroscopici irreversibili, come quelli legati a fenomeni di attrito, di viscosità, di diffusione, di conduzione del calore, ecc., le corrispondenti leggi fenomenologiche non sono invarianti per inversione temporale. Cosí per esempio l'attrito e la viscosità si oppongono sempre al moto dei corpi e al flusso dei fluidi e mai li agevolano, le sostanze in soluzione diffondono spontaneamente dalle zone a concentrazione piú alta verso quelle dove la concentrazione è piú bassa e mai viceversa, se si pongono a contatto corpi a temperatura diversa allora il calore passa spontaneamente dal corpo a temperatura piú alta verso quello a temperatura piú bassa, mentre non accade mai che il corpo a temperatura piú bassa si raffreddi ulteriormente cedendo spontaneamente calore a quello a temperatura piú alta. Pertanto una ben definita unidirezionalità temporale può essere associata a tutti i processi macroscopici che coinvolgono fenomeni dissipativi.

Questa unidirezionalità temporale è in ultima analisi connessa con una importante legge fisica, di validità generale, nota come legge di accrescimento dell'entropia, che può essere posta alla base del secondo principio della termodinamica. La conduzione del calore, i fenomeni di diffusione, i processi nei quali intervengono effetti di attrito e di viscosità ed in generale tutti i processi macroscopici irreversibili sono accompagnati da una generale crescita locale di entropia. Anche i processi che avvengono nel nostro organismo, in particolare quello di memorizzazione, sono associati a un inevitabile incremento dell'entropia.

Si ha quindi in definitiva che la legge di accrescimento dell'entropia, in quanto legge fisica generale che traduce l'irreversibilità dei processi macroscopici, induce una differenza qualitativa tra eventi passati ed eventi futuri, che viene da noi interpretata come flusso direzionale irreversibile del tempo stesso. Si viene cosí a riconoscere che in ultima istanza il secondo principio della termodinamica, con la sua legge di accrescimento dell'entropia, deve essere considerato alla base della nostra concezione psicologica ed empirica del carattere irreversibile del tempo e quindi del concetto stesso generale di tempo anche nelle sue implicazioni storiche e filosofiche.

È possibile quindi ottenere la sostanziale riduzione della dicotomia reversibilità/irreversibilità ad un ambito strettamente fisico, piú facilmente affrontabile con i metodi d'indagine scientifica attualmente disponibili. In questo ambito il problema può essere posto nei termini di individuare le forme e i

modi secondo cui è possibile conciliare sulla base di una interpretazione unitaria dei fenomeni fisici, il carattere temporalmente reversibile delle leggi elementari microscopiche del moto con le proprietà di irreversibilità delle leggi fenomenologiche dei processi macroscopici dissipativi.

Il problema cosí posto è storicamente riconducibile al programma di Ludwig Boltzmann [1896-98] di dare una fondazione microscopica al secondo principio della termodinamica, nel quadro di un'interpretazione cinetico-molecolare della struttura dei corpi materiali.

Nelle sue linee concettuali di base, la soluzione sostanziale di questo problema è stata ottenuta nel corso di un'intensa e multiforme attività di ricerca, che ha avuto i suoi momenti decisivi verso la fine del secolo scorso, all'epoca della polemica tra i sostenitori dell'ipotesi atomica, Boltzmann in testa, e i suoi detrattori, con l'individuazione di tutti gli elementi essenziali, di natura fisica e probabilistica, che sono alla base della possibilità di superamento della dicotomia e dei quali verrà data una trattazione per quanto possibile esauriente nelle pagine seguenti. La caratteristica determinante della soluzione della dicotomia, cosí come si è venuta storicamente configurando, è costituita dal riconoscimento che le equazioni fenomenologiche, che descrivono i processi macroscopici irreversibili, e piú in generale il secondo principio della termodinamica, non hanno validità assoluta e illimitata ma devono essere intesi come validi con altissima probabilità, non potendosi escludere, per via puramente dinamica, che possano esistere situazioni microscopiche patologiche che diano luogo a comportamenti macroscopici in disaccordo con le equazioni fenomenologiche e il secondo principio della termodinamica. Peraltro tali situazioni microscopiche patologiche hanno, per sistemi sufficientemente complessi, una trascurabile probabilità di verificarsi, per cui in tutti i casi pratici il secondo principio della termodinamica e le equazioni fenomenologiche risultano pienamente soddisfatti.

D'altra parte il problema è da considerare tuttora aperto, nel senso che non esiste ancora una derivazione completamente rigorosa delle equazioni fenomenologiche dei processi irreversibili a partire dalle leggi fondamentali del moto dei costituenti elementari della materia. Al termine di questo articolo si descriveranno le difficoltà che ancora si oppongono ad una completa realizzazione del programma e verrà data una rapida illustrazione di alcune delle linee di ricerca attualmente perseguite e volte ad una progressiva rimozione degli ostacoli tuttora esistenti.

Per ragioni di semplicità la trattazione sarà limitata ai casi coperti dalla meccanica classica (non quantistica), che permettono comunque di isolare nella forma piú efficace, sia dal punto di vista storico sia da quello concettuale, tutti i punti determinanti del problema della dicotomia reversibilità/irreversibilità e della sua soluzione su basi probabilistiche.

Occorre infine rilevare come le grandi difficoltà, sia di tipo fisico sia di tipo matematico, che impediscono tuttora una soluzione completamente soddisfacente e definitiva della dicotomia, abbiano indotto alcuni ricercatori e le loro scuole ad una drastica revisione dei fondamenti stessi del problema, volta a ricercare le basi dell'irreversibilità non nella meccanica (classica o quantistica)

reversibile ma in una nuova meccanica di tipo stocastico già intrinsecamente irreversibile.

A questo proposito può essere istruttivo confrontare le affermazioni secondo cui «l'irreversibilità... non è affatto in contraddizione con le leggi della dinamica; al contrario essa segue da queste leggi nei casi in cui venga raggiunto un sufficiente grado di "complessità"» [Prigogine e Nicolis 1977, p. 2] con le seguenti: «Bisogna dunque arrendersi all'evidenza: fino a che le premesse medesime – vale a dire la descrizione dinamica e la sua incarnazione in meccanica statistica, l'equazione di Liouville – non saranno modificate, l'entropia, come funzione crescente in maniera monotòna fino a un massimo, resterà sprovvista di significato dinamico. L'irreversibilità non può essere dedotta dalla dinamica; essa deve dunque esservi introdotta, se la dinamica vuole conservare una certa pertinenza per quello che concerne i processi costitutivi di questo mondo in divenire, la natura» [Prigogine e Stengers 1978, p. 432].

Tuttavia l'introduzione di dinamiche di tipo stocastico, intrinsecamente irreversibili e soggette a sorgenti esterne di disordine, non fa che spostare i termini del problema, tramite l'assunzione come postulati proprio delle caratteristiche essenziali di ciò che si vorrebbe spiegare (irreversibilità ed effetti dissipativi). Inoltre si produce un lacerante dualismo tra le proprietà del moto dei sistemi con un piccolo numero di costituenti, descritto senza dubbio dalla meccanica (classica o quantistica) reversibile, e le proprietà del moto dei sistemi termodinamici con un enorme numero di costituenti, descritto da un'ipotetica meccanica di tipo stocastico intrinsecamente irreversibile, con conseguenze imprevedibili per la stessa validità dell'ipotesi atomica. Le impostazioni di questo tipo quindi, nonostante il loro profondo interesse di natura puramente fenomenologica, non possono costituire la base per l'interpretazione di fenomeni complessi, quali l'irreversibilità e gli effetti dissipativi, tramite teorie piú semplici, quali la meccanica (classica o quantistica) reversibile.

Sembra quindi piú opportuno seguire l'impostazione complessiva del problema delineata precedentemente, che, a partire dalle profonde intuizioni di Boltzmann, ha avuto un progressivo fecondo arricchimento nel suo travagliato e difficile sviluppo storico, fornendo continuamente occasioni e spunti per nuove e piú approfondite investigazioni del piú alto interesse fisico e matematico.

Questo punto di vista sembra in particolare significativamente confortato da importanti risultati di indagini svolte recentemente con metodi matematici rigorosi sulle proprietà dinamiche di sistemi con un numero molto grande di di gradi di libertà, nel quadro del problema delle connessioni tra termodinamica e meccanica. Sulla base di queste indagini, di cui si darà un cenno nel seguito, e in notevole sintonia con una linea di pensiero, che per la sua continuità storica costituisce uno dei patrimoni piú importanti (e non a caso ripetutamente contestato) della cultura scientifica moderna, si è potuto affermare che «i nostri risultati... mostrano che il decremento della funzione H di Boltzmann è una conseguenza della meccanica classica (e della scelta opportuna delle condizioni iniziali) e che quindi non vi è necessità di introdurre sorgenti esterne di disordine... per produrre un comportamento irreversibile» [Lanford 1976, p. 133].

1. *Essenziale reversibilità delle leggi elementari del moto.*

Per poter illustrare il primo termine della dicotomia si prenderanno in esame alcuni degli aspetti generali della meccanica classica, rinviando ad altri articoli di questa stessa *Enciclopedia* per una piú completa trattazione delle leggi elementari del moto sia nel caso della meccanica, sia in quelli dell'elettromagnetismo, della meccanica quantistica, della relatività, ecc.

Nella meccanica classica lo stato di un sistema con n gradi di libertà è individuato, in un generico istante t, dalle sue n coordinate q_i, che ne descrivono la configurazione, e dai loro n momenti coniugati p_i, che rappresentano gli impulsi ($i=1, 2, \ldots, n$). L'insieme delle specificazioni (q_i, p_i) definisce un punto P in uno spazio delle fasi (spazio Γ) con $2n$ dimensioni. Si chiamano osservabili le funzioni generiche, $A(p, q)$, delle p e delle q. Il loro valore resta quindi univocamente determinato istante per istante se è noto lo stato, cioè il complesso dei valori delle p e delle q.

Se H è la funzione di Hamilton (energia) del sistema, allora il moto naturale del punto rappresentante lo stato nello spazio delle fasi deve soddisfare le equazioni di Hamilton

$$\dot{q}_i=\frac{\partial H}{\partial p_i}=\{q_i, H\} \qquad \dot{p}_i=-\frac{\partial H}{\partial q_i}=\{p_i, H\},$$

($i=1, 2, \ldots, n$), dove $\dot{q}_i$ e $\dot{p}_i$ sono le derivate temporali di $q_i(t)$ e $p_i(t)$ e $\{\ ,\ \}$ sono le parentesi di Poisson definite in generale per due osservabili $A(p, q)$ e $B(p, q)$ da

$$\{A, B\}=\sum_i\left(\frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i}-\frac{\partial A}{\partial p_i}\frac{\partial B}{\partial q_i}\right).$$

In generale, per una qualsiasi osservabile $A(p, q)$, dipendente dal tempo solo implicitamente attraverso la dipendenza delle p e delle q, si ha la seguente equazione di evoluzione temporale $\dot{A}=\{A, H\}$. In particolare, quindi, le osservabili che hanno parentesi di Poisson nulla con l'hamiltoniana sono delle costanti del moto.

Nei casi fisici che qui interessano la funzione di Hamilton $H=H(p, q)$ non dipende esplicitamente dal tempo (quindi l'energia è una costante del moto) ed assume la forma $H=T+U$, dove T è l'energia cinetica e U l'energia potenziale del sistema.

Si consideri ad esempio un sistema composto da N particelle puntiformi di massa m poste in un recipiente ed interagenti attraverso un potenziale centrale a due corpi, per cui $n=3N$; le $q_1, \ldots, q_n$ allora rappresentano le coordinate cartesiane lungo gli assi x, y, z di ciascuna delle particelle, le $p_1, \ldots, p_n$ le corrispondenti componenti degli impulsi, mentre si ha

$$T=\sum_{i=1}^{n}\frac{p_i^2}{2m} \qquad U=U(q)=\sum_{j=1}^{N}V_1(\mathbf{r}_j)+\sum_{j<j'}^{1\ldots N}V_2(|\mathbf{r}_j-\mathbf{r}_{j'}|),$$

dove $\mathbf{r}_j$ è la proiezione della particella j-esima (in particolare $\mathbf{r}_1 \equiv (r_{1x}, r_{1y}, r_{1z})$ con $r_{1x} = q_1$, $r_{1y} = q_2$, $r_{1z} = q_3$, ecc.) e V_1 il potenziale che schematizza l'interazione con il recipiente, mentre V_2 è il potenziale centrale a due corpi, che descrive l'interazione tra tutte le coppie di particelle j, j' con $j < j'$.

Le equazioni di Hamilton determinano univocamente la traiettoria del punto P nello spazio Γ: una volta assegnata la posizione iniziale $P = P(t)$ in un dato istante t resta univocamente determinata in linea di principio, in assenza di singolarità nelle equazioni, la posizione $P(t)$ in ciascun istante t (carattere deterministico dell'evoluzione).

Il concetto di reversibilità temporale può essere introdotto come segue. Si definisce il tempo invertito $t' = -t$ e la posizione invertita $P'(t')$ nello spazio Γ in modo che $q_i'(t') = q_i(t)$ e $p_1'(t') = -p_i(t)$. Poiché a causa della forma postulata per H si ha $H'(p', q') = H(p, q)$, si ottiene che per P' valgono ancora le equazioni di Hamilton

$$\dot{q}_i' = \frac{\partial H'}{\partial p_i'} \qquad \dot{p}_i' = -\frac{\partial H'}{\partial q_i'},$$

dove però al primo membro compaiono le derivate rispetto a t'. Si ottiene quindi il ben noto risultato che le leggi della meccanica sono completamente invarianti rispetto ad inversioni temporali. Se $P(t)$ al variare di t è una traiettoria naturale, lo è allora anche $P'(t')$, che corrisponde alla stessa traiettoria, ma percorsa in senso inverso. La meccanica pertanto non permette di stabilire una direzione privilegiata per il decorso del tempo, entrambe le direzioni sono completamente equivalenti e i due concetti di passato e futuro sono fra loro liberamente cambiabili.

Analoghe considerazioni valgono nel caso delle altre leggi elementari. Per esempio: la teoria elettromagnetica, descritta dalle equazioni di Maxwell nel vuoto

$$\begin{cases} \nabla \cdot \mathbf{E} = \varepsilon_0^{-1} \rho \\ \nabla \cdot \mathbf{B} = 0 \\ \nabla \times \mathbf{E} = -\dfrac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \mu_0 \varepsilon_0 \dfrac{\partial \mathbf{E}}{\partial t}, \end{cases}$$

è invariante sotto l'inversione temporale $t \to t' = -t$, purché per il campo elettrico $\mathbf{E}$, il campo magnetico $\mathbf{B}$, la densità di carica elettrica ρ e la densità di corrente elettrica $\mathbf{j}$, si assumano le seguenti leggi di trasformazione:

$$\begin{aligned} \mathbf{E} &\to \mathbf{E}'(\mathbf{x}, t') = \mathbf{E}(\mathbf{x}, t) \\ \mathbf{B} &\to \mathbf{B}'(\mathbf{x}, t') = -\mathbf{B}(\mathbf{x}, t) \\ \rho &\to \rho'(\mathbf{x}, t') = \rho(\mathbf{x}, t) \\ \mathbf{j} &\to \mathbf{j}'(\mathbf{x}, t') = -\mathbf{j}(\mathbf{x}, t). \end{aligned}$$

Naturalmente ε_0 e μ_0 sono delle costanti, dipendenti dal sistema di unità di misura scelto.

Anche per i sistemi quantistici, descritti dall'equazione di Schrödinger

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi,$$

dove ψ è la funzione d'onda e H l'operatore hamiltoniano, vale una proprietà di invarianza per inversione temporale, $t \to t' = -t$, purché si assuma $H \to H' = H$ (come nel caso classico trattato precedentemente) e inoltre $\psi \to \psi'(t') = \psi^*(t)$ (coniugazione complessa). In questo caso si vede immediatamente che per la ψ' vale ancora la citata equazione di Schrödinger. Quindi, anche a livello di meccanica quantistica, le due direzioni di decorso del tempo sono completamente equivalenti e non è possibile trascegliere una direzione privilegiata mediante considerazioni puramente interne all'ambito delle leggi dinamiche elementari.

2. *Le basi fenomenologiche della termodinamica degli stati di equilibrio e dei processi dissipativi irreversibili.*

Come si è visto all'inizio di questo articolo, dal punto di vista fisico, caratteri di irreversibilità temporale nella descrizione dei fenomeni naturali possono essere individuati nell'ambito di considerazioni termodinamiche.

La termodinamica è una teoria fenomenologica che descrive i processi macroscopici che coinvolgono produzione e trasmissione del calore.

Nei casi piú semplici i sistemi fisici sono descritti da un numero limitato di parametri interni, macroscopicamente misurabili, tra i quali sussistono determinate relazioni, ottenute per via sperimentale. In generale, per caratterizzare un sistema termodinamico è necessario specificare le sue proprietà fondamentali di tipo fisico e chimico, quali la massa totale, la composizione chimica, ecc., e descrivere poi le frontiere che lo separano dal mondo esterno (pareti di un contenitore, partizioni tra differenti parti del sistema, campi esterni, ecc.). Quindi un sistema termodinamico è caratterizzato da parametri interni ed esterni. I parametri che possono essere determinati mediante mezzi d'influenza esterna sul sistema, per esempio controllando corpi o campi esterni, sono chiamati parametri esterni. Invece i parametri interni sono determinati dallo stato del sistema stesso in corrispondenza di dati valori dei parametri esterni. Naturalmente la distinzione dei parametri macroscopici in interni ed esterni dipende dal modo particolare con cui il sistema viene separato dall'ambiente circostante. Un semplice esempio basta a chiarire questo punto fondamentale. Si consideri un gas contenuto in un cilindro verticale chiuso dall'alto mediante un pistone scorrevole a tenuta. Se il pistone viene fissato rigidamente al cilindro allora il volume occupato dal gas deve essere considerato come parametro esterno, mentre la pressione del gas è un parametro interno. Se invece il pistone è libero di scorrere ma è caricato mediante un peso definito, allora la pressione è un parametro esterno mentre il volume va considerato come un parametro interno.

Facendo riferimento ad altri articoli di questa stessa *Enciclopedia* per un trattamento piú approfondito, ecco sinteticamente le proposizioni fondamentali della termodinamica degli stati di equilibrio e di non-equilibrio.

2.1. Il postulato dell'esistenza di uno stato di equilibrio termodinamico.

L'evidenza accumulata nel corso delle piú svariate osservazioni empiriche conduce ad ammettere che per ogni sistema termodinamico, soggetto a determinate condizioni esterne, esiste uno stato di equilibrio termodinamico, nel quale i parametri macroscopici rimangono immutati nel corso del tempo e dal quale il sistema non può allontanarsi spontaneamente. Il concetto di equilibrio termodinamico è reso piú preciso dall'assunzione implicita che ogni sistema termodinamico fuori dall'equilibrio, lasciato soggetto a determinate condizioni esterne, tenderà a raggiungere nel tempo uno stato di equilibrio termodinamico. Si può qui osservare che già questo primo postulato fondamentale della termodinamica stabilisce una direzione privilegiata per il decorso del tempo, poiché stati di equilibrio sono spontaneamente raggiunti a partire da stati di non-equilibrio mentre il viceversa non è possibile.

2.2. La nozione di temperatura.

La temperatura viene introdotta come parametro caratterizzante gli stati di equilibrio, in modo tale che sia verificato il principio zero della termodinamica che può essere formulato come segue. Se due sistemi separatamente in equilibrio, isolati adiabaticamente dall'esterno (tali cioè che non possono scambiare calore con l'esterno), sono posti in contatto termico, allora i loro stati di equilibrio non sono alterati, e il sistema totale si trova in uno stato di equilibrio, solo se i sistemi iniziali hanno la stessa temperatura.

Una immediata conseguenza è che i parametri interni di un sistema all'equilibrio sono univocamente determinati dai valori dei parametri esterni e dalla temperatura.

2.3. Il primo principio della termodinamica.

Esprime il principio di conservazione dell'energia, in forma tale da permettere l'introduzione del concetto di quantità di calore per denotare l'energia trasferita ad un sistema senza compiere lavoro su di esso ma ponendolo in contatto termico con altri corpi. Esso può essere enunciato nella forma seguente. Per ogni sistema termodinamico esiste una funzione di stato U (energia interna), determinata a meno di una costante additiva arbitraria, tale che per ogni trasformazione infinitesima (cioè tra stati di equilibrio vicinissimi) si ha $dU = \delta L + \delta Q$, dove dU rappresenta la variazione di energia interna del sistema, δL il lavoro fatto dalle forze esterne durante la trasformazione e δQ il calore ceduto dall'esterno (misurato in unità opportune). Si noti che sono stati utilizzati simboli diversi per denotare le quantità infinitesime di lavoro e calore, δL e δQ, rispetto alla variazione infinitesima di energia interna dU, poiché non esistono

funzioni di stato di cui δL e δQ si possano considerare i differenziali. In generale δL e δQ dipendono separatamente dalla particolare trasformazione scelta per andare dallo stato iniziale a quello finale, mentre il primo principio assicura che la loro somma $\delta L + \delta Q$ non dipende dalla trasformazione poiché è il differenziale di una funzione di stato.

2.4. Il secondo principio della termodinamica.

Limita le possibilità di trasformazioni permesse dal primo principio. Esso può essere basato o sulla formulazione di Clausius, secondo cui non è possibile realizzare dispositivi in grado di trasferire calore da un corpo piú freddo a un corpo piú caldo senza che altri effetti siano prodotti, oppure sulla formulazione di Thomson (Lord Kelvin), secondo cui è impossibile costruire un perpetuum mobile di seconda specie, cioè un dispositivo che, funzionando in maniera ciclica, produca come unico effetto la trasformazione del calore estratto da una sola sorgente completamente in lavoro. Non è difficile mostrare l'equivalenza delle due formulazioni. Anche qui si osservi che, per esempio, la formulazione di Clausius, secondo cui non è possibile che il calore fluisca spontaneamente da un corpo piú freddo a uno piú caldo, introduce ancora una sostanziale irreversibilità nella descrizione temporale dei fenomeni.

Un'importante conseguenza del secondo principio è la possibilità d'introdurre una temperatura assoluta T e una nuova funzione di stato, l'entropia S, in modo tale che per le trasformazioni quasi-statiche, che cioè avvengono attraverso la successione di stati di equilibrio infinitesimamente vicini, si abbia $\delta Q = T dS$. Questo corrisponde a riconoscere la quantità T^{-1} come fattore integrante del calore δQ scambiato, tale cioè che anche se δQ in generale non è un differenziale esatto invece $T^{-1}\delta Q$ lo è. Se invece δQ è il calore scambiato in una trasformazione generica (non necessariamente quasi-statica) allora si ha $TdS \geq \geq \delta Q$. Il secondo principio assicura poi che l'entropia è sempre crescente nel corso di una trasformazione irreversibile da uno stato di equilibrio a un altro (di una trasformazione cioè che non procede attraverso la successione di stati di equilibrio, come per esempio il livellamento spontaneo di temperatura di due corpi inizialmente a temperature diverse e posti a contatto termico). Si ha quindi ancora una conferma della sostanziale irreversibilità temporale introdotta dal secondo principio della termodinamica.

2.5. Le equazioni cinetiche della termodinamica del non-equilibrio.

Allo scopo di fornire una base piú ampia per la discussione del superamento della dicotomia reversibilità/irreversibilità, specialmente in riferimento alle direzioni lungo le quali si svolge la ricerca contemporanea, è opportuno anche introdurre una sommaria descrizione delle equazioni fenomenologiche che regolano i processi irreversibili del non-equilibrio, quali per esempio la conduzione del calore, la diffusione, l'attrito viscoso, la conducibilità elettrica, ecc.

Per semplicità ci si limiterà a considerare il caso di un fluido isotropo viscoso

e conduttore del calore il cui stato fisico è caratterizzato, istante per istante, da una serie di parametri fenomenologici locali dipendenti dal posto, legati tra di loro dalle relazioni che verranno ora descritte.

Si considerino anzitutto la densità di massa del fluido $\rho(\mathbf{x}, t)$ e la densità di corrente $\mathbf{j}(\mathbf{x}, t)$ per le quali vale l'equazione di continuità

$$\frac{\partial\rho}{\partial t}+\nabla\cdot\mathbf{j}=0,$$

che esprime, in termini locali, il principio di conservazione della massa.

È anche conveniente introdurre il campo di velocità media di flusso $\mathbf{u}(\mathbf{x}, t)$, di componenti $u_i(\mathbf{x}, t)$, $i=1, 2, 3$, in modo che sia $\mathbf{j}(\mathbf{x},t)=\rho(\mathbf{x},t)\,\mathbf{u}(\mathbf{x},t)$, e il tensore degli sforzi $P_{ij}(\mathbf{x}, t)$, $i,j=1, 2, 3$. Allora il secondo principio della dinamica (conservazione dell'impulso) assume la forma

$$\rho\frac{Du_i}{Dt}=\rho a_i-\frac{\partial}{\partial x_j}P_{ij}$$

($i=1, 2, 3$), dove si è utilizzata la convenzione di somma sugli indici ripetuti e si è introdotta la derivata temporale sostanziale

$$\frac{D}{Dt}=\frac{\partial}{\partial t}+\mathbf{u}\cdot\nabla,$$

mentre $\mathbf{a}(\mathbf{x}, t)$ è la forza esterna per unità di massa agente sugli elementi di fluido intorno al punto $\mathbf{x}$ all'istante t. Il termine che coinvolge il tensore P_{ij} nel secondo principio della dinamica è connesso con le forze interne scambiate tra i vari elementi del fluido a causa degli effetti di pressione e di viscosità.

Infine per tener conto degli scambi energetici che avvengono all'interno del fluido conviene introdurre la densità di energia interna $\varepsilon(\mathbf{x}, t)$, la densità del flusso di corrente di calore $\mathbf{q}(\mathbf{x}, t)$ e il tensore della velocità di deformazione

$$D_{ij}=\frac{1}{2}\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)$$

($i,j=1, 2, 3$). Allora il primo principio della termodinamica assume la forma locale

$$\rho\frac{D}{Dt}\left(\frac{\varepsilon}{\rho}\right)+\nabla\cdot\mathbf{q}=-P_{ij}\,D_{ij},$$

dove i tre termini sono connessi rispettivamente con la variazione di energia interna di un volumetto infinitesimo, il calore scambiato con l'esterno e il lavoro fatto dalle forze di pressione e di viscosità.

La legge empirica di Newton per la viscosità permette immediatamente di connettere il tensore degli sforzi P_{ij} con la pressione $p(\mathbf{x}, t)$ e il tensore della velocità di deformazione D_{ij} tramite una relazione della forma

$$P_{ij}=p\delta_{ij}-2\eta\left(D_{ij}-\frac{1}{3}D_{\kappa\kappa}\delta_{ij}\right)-\zeta D_{\kappa\kappa}\delta_{ij},$$

dove η e ζ sono i coefficienti di viscosità.

Invece la legge di Fourier sulla conduzione del calore permette di scrivere $\mathbf{q}=-\kappa\nabla T$, dove κ è il coefficiente di conducibilità termica e T la temperatura.

Si possono allora assumere come variabili indipendenti le cinque grandezze fisiche, variabili in generale da punto a punto all'interno del volume occupato dal fluido e nel tempo, $\rho(\mathbf{x}, t)$, densità di massa, $u_i(\mathbf{x}, t)$, componenti del campo di velocità media di flusso, e $T(\mathbf{x}, t)$, temperatura locale. Tutte le altre grandezze introdotte possono essere espresse mediante queste se si prendono in considerazione le leggi di Newton e di Fourier ed inoltre si tiene conto del fatto che sulla base dei principî della termodinamica è possibile esprimere la pressione p e la densità di energia interna ε in funzione della densità ρ e della temperatura T, in ogni punto del fluido.

Ma allora l'equazione di continuità, il secondo principio della dinamica e il primo principio della termodinamica, espressi nella forma locale introdotta precedentemente, costituiscono un sistema di cinque equazioni differenziali del primo ordine rispetto al tempo per le cinque variabili fondamentali ρ, $\mathbf{u}$ e T. Si ottiene quindi in conclusione che lo stato fisico del sistema può essere descritto istante per istante dai campi ρ, $\mathbf{u}$ e T, per i quali sussistono equazioni che permettono di determinarli, almeno in linea di principio, ad ogni successivo istante t, una volta che sia stato specificato lo stato iniziale a $t=0$ caratterizzato da $\rho(\mathbf{x}, 0)$, $\mathbf{u}(\mathbf{x}, 0)$ e $T(\mathbf{x}, 0)$.

Considerazioni analoghe valgono anche nei casi piú complessi. In generale l'evoluzione data dalle equazioni fenomenologiche dei processi macroscopici è completamente deterministica, come quella hamiltoniana dei sistemi meccanici, ma esiste una profonda differenza, costituita dalla sostanziale irreversibilità temporale, in contrasto con il caso hamiltoniano.

Per illustrare le irreversibilità nel caso considerato basta osservare che sotto l'inversione $t\to t'=-t$ le grandezze fondamentali dovrebbero trasformarsi secondo le relazioni

$$\begin{aligned}\rho&\to\rho'(\mathbf{x}, t')=\rho(\mathbf{x}, t)\\ \mathbf{u}&\to\mathbf{u}'(\mathbf{x}, t')=-\mathbf{u}(\mathbf{x}, t)\\ T&\to T'(\mathbf{x}, t')=T(\mathbf{x}, t),\end{aligned}$$

ma si vede immediatamente che, nel caso in cui almeno uno dei coefficienti η, ζ e κ è diverso da zero, le equazioni di evoluzione non sono invarianti, anzi si trasformano in equazioni assurde dal punto di vista fisico in cui le forze di viscosità agevolano il flusso del fluido invece di ostacolarlo e il calore fluisce spontaneamente dalle zone a temperatura piú bassa a quelle a temperatura piú alta. L'origine fisica dell'irreversibilità è quindi da vedersi proprio nella forma delle leggi empiriche di Newton sulla viscosità e di Fourier sulla conduzione del calore.

Si ottiene quindi in conclusione che nel caso generale le equazioni fenomenologiche dei processi macroscopici, che si sviluppano nei sistemi termodinamici fuori dall'equilibrio e che coinvolgono effetti dissipativi, quali la viscosità e la conduzione del calore, a causa della loro sostanziale irreversibilità temporale introducono una rigida differenziazione tra le due direzioni di decorso del tem-

po. Data la loro importanza nello svolgimento di tutti i processi macroscopici, esse sono in ultima istanza alla base della concezione di unidirezionale irreversibilità del concetto di tempo.

3. *L'ipotesi molecolare della materia e l'emergenza della dicotomia reversibilità/irreversibilità.*

Lo schema concettuale, entro il quale può essere affrontato il contrasto tra le proprietà di reversibilità temporale delle leggi elementari del moto e le proprietà di irreversibilità delle leggi fenomenologiche dei processi macroscopici, è costituito dall'ipotesi molecolare della materia. Secondo questa ipotesi la materia non è continua ed infinitamente divisibile, ma è composta da un grande numero di piccoli corpi che si chiameranno in generale molecole. A seconda della composizione e della natura del particolare corpo materiale preso in considerazione, le molecole possono essere di un solo tipo oppure di parecchi tipi. In questo contesto si possono considerare gli atomi liberi, gli ioni e gli elettroni semplicemente come tipi speciali di molecole. Le molecole individuali sono in generale troppo piccole per essere viste separatamente anche con l'ausilio del piú potente microscopio. Il dispiegarsi congiunto di ricerche sperimentali e teoriche ha condotto a precise ipotesi sulla struttura e l'interazione delle varie molecole.

L'ipotesi molecolare è di grande importanza in chimica e in fisica. In alcuni casi, ad esempio per alcuni problemi di cristallografia, è sufficiente considerare le molecole staticamente, ma in generale è necessario tener conto dei loro movimenti che, anche se non visibili individualmente, sono estremamente rapidi. Un'estensione importante dell'ipotesi molecolare è la cosiddetta teoria cinetica del calore, nella quale si assume che le molecole si muovano piú o meno rapidamente a seconda che il corpo di cui fanno parte sia piú caldo o piú freddo. L'energia interna di un corpo materiale viene allora interpretata come energia meccanica, cinetica e potenziale, dei movimenti molecolari.

L'ipotesi molecolare e la conseguente teoria cinetica del calore hanno dovuto combattere una lunga battaglia prima di trovare universale affermazione. Mentre le origini remote dell'ipotesi molecolare possono farsi risalire alle formulazioni filosofiche degli antichi atomisti (Leucippo, Democrito, Epicuro, Lucrezio), i primi tentativi d'interpretare i fenomeni termici in termini di movimenti molecolari furono fatti da Boyle e Newton. In quel tempo non si era ancora pervenuti ad una consistente teoria cinetica del calore e la teoria del flogisto sembrava piú soddisfacente, anche se non completamente sviluppata. Il flogisto era una sostanza ipotetica che si pensava essere una componente essenziale di tutti i corpi in grado di bruciare e dei metalli che era possibile ossidare, la quale veniva persa nel corso della combustione e dell'ossidazione. La teoria del flogisto fu formulata all'inizio del XVIII secolo ed ottenne largo riconoscimento, specialmente nel campo della chimica, finché Lavoisier non stabilí la vera natura della combustione negli anni verso la fine del secolo.

Dopo i primi tentativi di Boyle e Newton, la teoria del calore ebbe sviluppi importanti attraverso il lavoro di Daniele Bernoulli, che nel 1738 presentò il primo schema generale di una teoria cinetica dei gas svolta su basi quantitative e in grado di spiegare, in particolare, la legge di Boyle-Mariotte. Ma per piú di un secolo la teoria del calore dominante presso la maggioranza degli studiosi fu ancora costituita dalla teoria fenomenologica del flogisto, che raggiunse il maggiore successo con la scoperta del principio di Sadi Carnot del 1824, equivalente al secondo principio della termodinamica. In seguito i lavori di Joule, Mayer e Helmholtz mostrarono l'inconsistenza delle basi della teoria del flogisto e fecero pervenire alla formulazione corretta del primo principio della termodinamica come legge generale di conservazione dell'energia. Tuttavia la teoria del calore continuò a svilupparsi come teoria fenomenologica, anche se era ormai chiaro che fenomeni di tipo cinetico-molecolare ne dovevano essere alla base. L'ultimo trionfo della teoria fenomenologica fu la costruzione dello schema generale della termodinamica classica, ottenuto verso la metà del XIX secolo attraverso il lavoro di Clausius e Thomson.

La fase moderna della teoria cinetico-molecolare del calore ebbe inizio nel 1858, quando Clausius introdusse il concetto di cammino libero medio e riformulò i principî fondamentali dell'intera teoria, che poi vennero ulteriormente sviluppati in maniera decisiva da Maxwell e da Boltzmann.

Tuttavia per piú di mezzo secolo l'unica teoria accettata fu quella della termodinamica fenomenologica, i cui principî furono assunti come base di una interpretazione generale di natura puramente descrittiva e formale della fisica, che doveva escludere ogni considerazione ipotetica sulla struttura profonda dei fenomeni naturali e che quindi rigettava in linea di principio l'ipotesi molecolare e si affidava quasi esclusivamente a considerazioni di tipo energetico. Di conseguenza Boltzmann, che aveva dato una profonda formulazione della teoria del calore su basi cinetico-molecolari della quale si discuterà in seguito, fu costretto a combattere per tutta la sua vita una continua e dura battaglia contro gli energetisti, che contavano tra le loro file scienziati dal valore indiscusso, quali Ostwald e Mach.

Tuttavia alla fine le idee basate sull'ipotesi molecolare prevalsero, trovando conferma della loro correttezza nel 1905 attraverso i lavori di Einstein e Smoluchowski sul moto browniano, i cui risultati furono in seguito verificati sperimentalmente da Perrin.

Da allora la teoria cinetico-molecolare del calore ottenne il generale riconoscimento, anche grazie all'opera di Gibbs che pose le basi per la sua estensione e generalizzazione nella forma moderna di meccanica statistica.

Dopo questa breve parentesi storica bisogna ora tornare all'esposizione delle conseguenze della teoria cinetico-molecolare, mostrando in particolare la necessità dell'introduzione di considerazioni di natura statistica legate alla teoria della probabilità.

Per ottenere una maggiore semplicità di trattazione, si considererà nel seguito esclusivamente il caso di un gas monoatomico a bassa densità. In accordo con l'ipotesi molecolare, il sistema fisico costituito dal gas viene schematizzato

come l'insieme di un numero molto grande di corpi materiali puntiformi interagenti attraverso un potenziale a due corpi di forma adeguata. Per un tale sistema meccanico le considerazioni svolte precedentemente possono immediatamente applicarsi. Lo stato istantaneo del sistema sarà determinato da un punto nello spazio delle fasi che si muove in accordo con le equazioni reversibili del moto illustrate precedentemente. Nella meccanica ordinaria (per esempio nello studio del moto dei pianeti o dei satelliti artificiali) ci si pone generalmente il problema di determinare, sulla base delle note equazioni del moto, lo stato istante per istante una volta dato lo stato iniziale. Nel caso dei sistemi meccanici che costituiscono, nell'ambito dell'ipotesi cinetico-molecolare, dei modelli per i sistemi termodinamici questo modo di procedere non è possibile, per una varietà di ragioni, e metodi completamente diversi devono essere posti in uso. Infatti, se si considera per esempio un litro di gas elio a temperatura e pressione ordinaria, esso contiene ventisettemila miliardi di miliardi di atomi, un numero cioè incredibilmente alto che rende impossibile una trattazione puramente meccanica. Anzitutto sarebbe impossibile conoscere, in seguito a misure effettuate, le posizioni e le velocità di tutte le molecole in un dato istante iniziale (cioè lo stato meccanico iniziale). Inoltre, anche conoscendo lo stato iniziale, sarebbe impossibile risolvere, con ogni mezzo noto, le equazioni del moto di Hamilton per conoscere lo stato negli istanti successivi e comunque tale soluzione accurata sarebbe completamente inutile perché comunque lo stato iniziale sarebbe noto a meno di inevitabili errori sperimentali, che si amplificherebbero nel corso del moto, a causa dell'enorme complessità del sistema, fino a rendere inutilizzabili i risultati trovati per lo stato del sistema a un istante successivo.

Quindi è completamente inutile seguire l'andamento del moto delle singole molecole, ma conviene seguire una strada completamente diversa e concentrarsi sulla determinazione di proprietà statistiche, come per esempio i valori medi del numero, del momento e dell'energia delle molecole entro un elemento di volume, eventualmente mediati anche su piccoli intervalli temporali, oppure la distribuzione media della velocità o di altre caratteristiche dinamiche per queste molecole.

D'altra parte non sono solo ragioni di tipo matematico, legate all'enorme numero di componenti elementari e all'intrattabile complessità delle equazioni del moto, a forzare necessariamente l'introduzione di considerazioni di natura statistica nel trattamento dei sistemi termodinamici mediante l'ipotesi molecolare. In realtà esistono profonde ragioni di natura fisica che mostrano la completa adeguatezza dei metodi statistici. Occorre infatti tener conto del fatto che la descrizione macroscopica dei fenomeni che si svolgono in un sistema termodinamico, anche fuori dall'equilibrio, coinvolge solo i parametri locali effettivamente misurabili quali la densità di massa, la densità di corrente, il flusso di calore, la densità di energia interna, ecc., come è stato rapidamente illustrato precedentemente nel § 2. Nel passaggio quindi dalla descrizione microscopica in termini meccanici alla descrizione macroscopica basata sulle equazioni cinetiche di tipo fenomenologico, avviene una sorta di contrazione nelle variabili

essenziali da prendersi in considerazione per cui i parametri macroscopici fenomenologici devono considerarsi come ottenuti dai parametri microscopici tramite opportune medie di tipo sia spaziale sia temporale, come verrà mostrato piú dettagliatamente nel seguito, almeno in alcuni casi semplici.

È questa in definitiva la ragione piú profonda per l'introduzione di metodi statistici nella trattazione dei fenomeni termodinamici nell'ambito dell'ipotesi cinetico-molecolare. Nel corso di questi tentativi occorre considerare per esempio non solo la dinamica degli urti molecolari, ma anche la loro statistica, utilizzando opportune assunzioni probabilistiche. I fondatori della moderna teoria cinetico-molecolare utilizzarono dapprima le necessarie considerazioni di natura probabilistica basandosi su argomenti di carattere intuitivo, spesso solo implicitamente utilizzati. Le loro ricerche stabilirono i fondamenti di una estesa disciplina, parte della fisica teorica, nota come meccanica statistica, in grado di trattare, almeno in linea di principio, tutti i sistemi termodinamici all'equilibrio e fuori dell'equilibrio.

I metodi della meccanica statistica, specialmente per i fenomeni del non-equilibrio, rivestono attualmente un grande interesse, non solo per gli stimolanti problemi aperti nell'ambito della critica sui fondamenti, ma anche per le loro possibilità di applicazione a larghi campi della scienza e della tecnologia attuali, quali la fisica dei plasmi, di immediato interesse per la soluzione del problema della fusione termonucleare controllata, la gas-dinamica, la tecnologia spaziale, la conversione dell'energia, l'ingegneria chimica, la fisica atmosferica, la fisica dello stato solido, ecc.

La meccanica statistica applica metodi di natura probabilistica a problemi di tipo meccanico e, nonostante i suoi innegabili successi, i suoi principî fondamentali non sono privi di ambiguità e di oscurità. Queste difficoltà furono già avvertite dai fondatori e costituirono la base di tutte le polemiche, che peraltro contribuirono moltissimo a una continua chiarificazione dei fondamenti. Chiarificazione che è continuata negli sviluppi futuri anche se la teoria non ha ancora raggiunto, specialmente per quanto riguarda i fenomeni del non-equilibrio, una formulazione definitiva e completamente consistente.

È chiaro che una delle maggiori sorgenti di difficoltà è costituita proprio dall'apparente contraddizione tra la reversibilità temporale delle equazioni della meccanica, poste a base dell'interpretazione cinetico-molecolare, e la sostanziale irreversibilità insita nel secondo principio della termodinamica e nella descrizione fenomenologica dei processi del non-equilibrio, che sono proprio le caratteristiche macroscopiche dei sistemi termodinamici che la teoria cinetica dovrebbe spiegare. Rinviando ai successivi paragrafi per una trattazione piú approfondita, ci si limita qui ad anticipare che le considerazioni di natura probabilistica svolgono un ruolo determinante nella soluzione di questo paradosso, nel senso che le affermazioni fenomenologiche della termodinamica (per esempio la validità delle equazioni cinetiche) devono essere intese non come aventi validità assoluta ma solo come altamente probabili, non potendosi escludere, su basi puramente meccaniche, che particolari stati microscopici, appositamente costruiti ed ai quali corrispondono ragionevoli caratterizzazioni macroscopiche,

possano evolvere in maniera tale che le corrispondenti variabili macroscopiche mostrino un comportamento deviante rispetto alle leggi della termodinamica. È compito della teoria mostrare come e in che senso tali stati microscopici patologici abbiano scarsa probabilità di verificarsi. È questo un punto tuttora aperto sul quale ci si soffermerà a lungo nel seguito.

4. *Spiegazione delle leggi della termodinamica dei sistemi in equilibrio attraverso i metodi della meccanica statistica.*

Dopo aver introdotto l'ipotesi molecolare, occorre mostrare brevemente come sia possibile pervenire a una spiegazione completa delle leggi della termodinamica dei sistemi in equilibrio mediante una serie di prescrizioni di natura meccanica e probabilistica, i cui fondamenti furono posti da Maxwell, Boltzmann e Gibbs.

È anzitutto opportuna l'introduzione del concetto di insieme statistico.

Si consideri un sistema meccanico del tipo trattato nel § 1. Con il trascorrere del tempo il punto $P(t)$ rappresentativo del sistema descrive un'orbita nello spazio Γ. Le coordinate q_i cambiano di valore a causa delle velocità delle particelle, mentre le p_i evolvono a causa dell'azione delle forze interne ed esterne che agiscono sulle particelle. Poiché si è supposto che la funzione di Hamilton non dipenda esplicitamente dal tempo, si ha che l'energia totale $\varepsilon(p, q)$ è una costante del moto. L'orbita del punto rappresentativo sarà quindi situata su una ipersuperficie Σ (chiamata superficie dell'energia) di equazioni $\varepsilon(p, q) = c$, dove la costante c dipende naturalmente dallo stato iniziale del sistema.

Invece di considerare un solo sistema è spesso vantaggioso, o come si mostrerà in seguito addirittura inevitabile, considerare un grande numero di sistemi simili completamente indipendenti, ciascuno dei quali contenga ancora N particelle che si muovono con le stesse equazioni del moto in identiche circostanze fisiche. In altri termini, tutti questi sistemi avranno la stessa hamiltoniana e lo stesso spazio delle fasi Γ, ma i loro punti rappresentativi non saranno in generale gli stessi. Si supponga di considerare un numero s elevatissimo di tali sistemi in modo che sia lecito introdurre una densità $D(p, q, t)$, con la quale i punti rappresentativi sono distribuiti sullo spazio delle fasi. Una tale collezione di sistemi simili e indipendenti venne chiamata da J. Willard Gibbs [1902] insieme statistico.

È conveniente introdurre nello spazio delle fasi Γ l'elemento di volume $d\Omega$ corrispondente a coordinate negli intervalli $(p_i, p_i + dp_i)$ e $(q_i, q_i + dq_i)$, in modo da avere $d\Omega = \prod_i dp_i dq_i$.

Sulla base delle equazioni di Hamilton si può immediatamente dimostrare che se $\{p', q'\}$ sono le coordinate del punto rappresentativo all'istante t', per una traiettoria che passa per $\{p, q\}$ all'istante t, allora si ha $d\Omega = d\Omega'$, cioè l'estensione in fase di un volumetto $d\Omega$ si conserva nel tempo (teorema di Liouville). Questo risultato mostra l'utilità dell'introduzione del concetto di estensione in fase $d\Omega$.

Poiché i punti rappresentativi si muovono nello spazio delle fasi, è ovvio che la densità in fase $D(p, q, t)$ sarà in generale una funzione del tempo. Non è difficile dimostrare sulla base del teorema di Liouville che la variazione nel tempo di D in un certo punto dello spazio delle fasi è data da

$$\frac{\partial D}{\partial t} = \{H, D\}.$$

Ovviamente si deve avere anche $\int_\Gamma D(p, q, t)\, d\Omega = s$, dove l'integrazione è estesa sull'intera parte accessibile dello spazio delle fasi.

Un insieme per il quale $\partial D/\partial t = 0$ è chiamato stazionario e nel seguito di questo paragrafo ci si limiterà principalmente a considerare tali tipi di insiemi.

Naturalmente un insieme è stazionario se $\{H, D\} = 0$. Quindi in particolare saranno stazionari tutti gli insiemi in cui D è una assegnata funzione dell'energia $D = D(\varepsilon)$. Gli insiemi stazionari di questo tipo svolgono un ruolo molto importante dal punto di vista fisico per ragioni che saranno spiegate nel seguito in connessione con le discussioni sull'ipotesi ergodica.

È opportuno introdurre anche la densità dell'insieme (o coefficiente di probabilità) $\rho(p, q, t) = s^{-1} D(p, q, t)$, normalizzato cosí che $\int_\Gamma \rho(p, q; t)\, d\Omega = 1$.

Seguendo Gibbs va introdotto anche l'indice di probabilità $\eta(p, q; t)$ definito da $\eta = \log \rho$. Sulla base delle assunzioni fatte ci si limiterà a considerare casi in cui $\rho = \rho(\varepsilon)$, $\eta = \eta(\varepsilon)$.

Una funzione $G(p, q)$ dipendente dalle p_i e q_i è detta una funzione di fase e il suo valore medio preso sull'insieme statistico ρ è dato da $\bar{G} = \int_\Gamma G\rho\, d\Omega$. Naturalmente, se l'insieme ρ è stazionario allora $\bar{G}$ è una costante indipendente dal tempo.

Occorre ora introdurre due tipi di insiemi stazionari, il canonico e il microcanonico, che verranno utilizzati nella successiva discussione. Ci si limita naturalmente a insiemi con densità dipendenti dalla sola energia. Si osservi a questo punto che in generale l'energia sarà una funzione delle p e delle q del sistema, come è stato spiegato nel § 1, che in piú però dipenderà da parametri esterni, a_i, come per esempio il volume occupato dal sistema oppure l'intensità dei campi esterni con cui il sistema interagisce.

Seguendo Gibbs va ora introdotto l'insieme canonico definito da

$$\eta = \frac{\psi - \varepsilon}{\vartheta} \qquad \rho = e^{(\psi - \varepsilon)/\vartheta},$$

dove ψ e ϑ sono delle costanti. La grandezza ϑ è detta modulo della distribuzione ed è connessa alla temperatura, come sarà dimostrato nel seguito. La grandezza ψ non dipende dalle p e q ma dipende da ϑ e dalle a_i, a causa della condizione di normalizzazione

$$\int_\Gamma \rho\, d\Omega = 1 \Rightarrow e^{-\psi/\vartheta} = \int_\Gamma e^{-\varepsilon/\vartheta}\, d\Omega.$$

L'insieme microcanonico è invece definito assumendo una distribuzione uniforme per la ρ nella regione dello spazio delle fasi contenuta tra le due su-

perfici di energia ε_0 ed $\varepsilon_0+\Delta$, e ponendo $\rho=0$ sulla parte restante dello spazio delle fasi. Nel caso in cui si faccia tendere Δ a zero allora l'insieme microcanonico è interamente concentrato sulla superficie dell'energia $\varepsilon(p, q)=\varepsilon_0$, sulla quale è distribuito con densità superficiale uniforme rispetto ad una misura $d\mu$, invariante nel tempo e normalizzata (misura microcanonica), la cui esistenza è una semplice conseguenza del teorema di Liouville precedentemente richiamato.

A questo punto è possibile spiegare per quale ragione vengono introdotti gli insiemi statistici dal momento che in realtà, nei casi fisici, esiste un unico sistema in considerazione, illustrando quindi la relazione tra gli insiemi statistici e i sistemi fisici realmente osservati.

La ragione è che il comportamento termodinamico all'equilibrio dei sistemi fisici può essere utilmente descritto mediante gli insiemi statistici. Inoltre, nel limite di volume infinito, gli insiemi statistici canonico e microcanonico, e gli altri quali l'insieme grancanonico pure introdotto da Gibbs, sono sostanzialmente equivalenti dal punto di vista fisico, a causa dello smorzarsi delle fluttuazioni nel limite in cui il numero delle particelle diviene infinito.

Un primo risultato, dovuto a Gibbs, mostra da solo l'importanza delle considerazioni di tipo probabilistico offerte dalla meccanica statistica nello studio della spiegazione della termodinamica dell'equilibrio.

Il comportamento macroscopico di un sistema fisico in equilibrio termico con un termostato a temperatura T può essere descritto correttamente, nell'ambito di una ben definita serie di prescrizioni, mediante il comportamento medio di sistemi che si trovano distribuiti secondo un insieme canonico caratterizzato da un modulo ϑ legato alla temperatura T dalla relazione $\vartheta=KT$, dove K è la costante di Boltzmann.

Data l'importanza di tale risultato se ne darà una dimostrazione dettagliata.

S'introducano anzitutto le forze generalizzate

$$A_i=-\frac{\partial\varepsilon}{\partial a_i},$$

dove $\varepsilon(p, q; \boldsymbol{a})$ è l'energia del sistema e $\{a_i\}$ sono i parametri esterni. Se le a_i sono soggette a variazioni da_i in modo da assumere i nuovi valori a_i+da_i, allora la variazione $d\varepsilon$ dell'energia sarà data da $d\varepsilon=-\sum_i A_i\,da_i$. Si consideri ora un nuovo insieme canonico corrispondente a nuovi valori del modulo $\vartheta+d\vartheta$ e dei parametri esterni a_i+da_i. Dalla condizione di normalizzazione $e^{-\psi/\vartheta}=\int_\Gamma e^{-\varepsilon/\vartheta}d\Omega$, tramite variazione rispetto a ϑ e a_i, si ottiene

$$-e^{-\psi/\vartheta}\,d\left(\frac{\psi}{\vartheta}\right)=\int_\Gamma e^{-\varepsilon/\vartheta}\left(-\varepsilon\,d\frac{1}{\vartheta}-\frac{1}{\vartheta}\,d\varepsilon\right)d\Omega,$$

da cui

$$d\left(\frac{\psi}{\vartheta}\right)=\bar\varepsilon\,d\left(\frac{1}{\vartheta}\right)+\left(\frac{1}{\vartheta}\right)\overline{d\varepsilon},$$

avendo pertanto sfruttato la definizione delle medie sull'insieme statistico fornita dalla formula seguente:

$$\bar G=e^{\psi/\vartheta}\int_\Gamma e^{-\varepsilon/\vartheta}G\,d\Omega=\int_\Gamma\rho G\,d\Omega.$$

Poiché si ha anche

$$d\frac{\bar\varepsilon}{\vartheta}=\frac{1}{\vartheta}\,d\bar\varepsilon+\bar\varepsilon\,d\frac{1}{\vartheta},$$

si ottiene in definitiva

$$d\frac{\bar\varepsilon-\psi}{\vartheta}=-d\bar\eta=\frac{d\bar\varepsilon-\overline{d\varepsilon}}{\vartheta},$$

che per analogia con il caso termodinamico descritto nel § 2 permette la seguente interpretazione. Anzitutto $d\bar\varepsilon$ può essere identificato con la variazione di energia interna dU, mentre $\overline{d\varepsilon}$ corrisponde al lavoro fatto dalle forze esterne quando i parametri esterni a_i variano da a_i ad a_i+da_i, quindi $\delta L=\overline{d\varepsilon}$ (notare che δL non è in generale un differenziale esatto poiché corrisponde alla media statistica della variazione di energia $d\varepsilon$ dovuta alla variazione dei parametri a_i). Si ha quindi anche l'identificazione di $d\bar\varepsilon-\overline{d\varepsilon}$ come calore δQ ceduto al sistema, in modo che il primo principio della termodinamica è soddisfatto nella forma

$$dU=\delta L+\delta Q,$$
$$d\bar\varepsilon=\overline{d\varepsilon}+(d\bar\varepsilon-\overline{d\varepsilon}).$$

Inoltre la relazione trovata precedentemente mostra che $1/\vartheta$ è un fattore integrante per δQ, quindi ϑ deve essere proporzionale alla temperatura assoluta, $\vartheta=KT$, e che inoltre $-\bar\eta$, a parte una costante additiva, deve essere identificata con l'entropia divisa per K, in modo che si abbia $dS=\delta Q/T$ a partire da $-d\bar\eta=(d\bar\varepsilon-\overline{d\varepsilon})/\vartheta$ e quindi $S=-K\bar\eta+\text{cost}$.

A rinforzare l'idea che l'insieme canonico possa essere utilizzato per descrivere un sistema fisico all'equilibrio termico si possono portare i seguenti ulteriori risultati dovuti ancora a Gibbs. Si può infatti dimostrare che se due insiemi canonici vengono accoppiati fisicamente, allora essi formeranno un insieme stazionario, ancora canonico, se i loro moduli sono uguali, se invece i loro moduli sono differenti allora in media l'insieme con modulo piú elevato perderà energia a favore di quello con modulo piú basso. Tutto questo risulta in accordo con le leggi fenomenologiche della termodinamica.

Per quanto riguarda l'insieme microcanonico, che è distribuito uniformemente su una superficie ad energia costante, ci si limita per il momento ad osservare che nel caso di sistemi con un numero molto grande di particelle, i risultati che si ottengono effettuando le medie rispetto all'insieme microcanonico sono gli stessi di quelli ottenuti tramite un insieme canonico opportunamente scelto, purché si trascurino effetti di fluttuazione nei valori delle variabili macroscopiche che comunque si annullano nel limite in cui il numero delle particelle diventa infinito.

Questa sostanziale equivalenza fisica, nel limite di sistemi infiniti, tra vari insiemi termodinamici, già individuata da Gibbs e progressivamente sviluppata in forma matematicamente rigorosa nella moderna meccanica statistica [cfr.

Ruelle 1969], mostra chiaramente l'enorme importanza del metodo basato sugli insiemi statistici per la spiegazione in termini meccanici e probabilistici del comportamento dei sistemi termodinamici all'equilibrio. Altri successi notevoli della teoria sono costituiti dalla derivazione di equazioni di stato a partire dalle interazioni fondamentali tra le particelle e dalla possibilità d'incorporare i fenomeni di transizione di fase nell'ambito di una prescrizione unitaria per la termodinamica dell'equilibrio.

Tuttavia è chiaro che il metodo basato sugli insiemi statistici, pur essendo di grande semplicità e potenza, ha un carattere di prescrizione ad hoc, lasciando completamente aperto il problema di determinare come e perché i sistemi termodinamici tendano all'equilibrio e perché l'equilibrio può essere descritto mediante gli insiemi termodinamici introdotti precedentemente. A questo scopo sono indispensabili metodi piú approfonditi basati su dirette considerazioni di natura dinamica e probabilistica, che permettono di dare un trattamento unificato e contemporaneo della meccanica statistica dell'equilibrio e del non-equilibrio.

Si darà nel seguito una rapida esposizione di questi metodi, partendo dai casi particolari trattati originariamente da Boltzmann; questo paragrafo va però terminato con un cenno schematico ai tentativi fatti per giustificare l'approccio all'equilibrio, descritto dagli insiemi statistici, mediante una costellazione di idee che va sotto il nome di teoria ergodica, di grande importanza sia per il suo interesse fisico e matematico agli effetti della spiegazione dei fondamenti della meccanica statistica, sia per le ripercussioni sul problema della definitiva soluzione della dicotomia reversibilità/irreversibilità, come sarà illustrato piú avanti. Si consideri un sistema meccanico isolato con una energia ε determinata. Il suo punto rappresentativo nello spazio Γ descriverà un'orbita sulla superficie dell'energia $\varepsilon(p, q) = \text{cost}$. L'orbita è completamente determinata dalle equazioni del moto, in particolare in ogni punto della superficie dell'energia la direzione lungo la quale si muoverà il punto è completamente determinata dai valori delle coordinate del punto stesso e quindi due orbite diverse non possono mai intersecarsi.

Boltzmann osservò che se un sistema isolato parte da una condizione termodinamica macroscopica di non-equilibrio, esso, sulla base dei principî della termodinamica, raggiungerà piú o meno velocemente una condizione di equilibrio e vi permarrà indefinitamente. Pertanto le medie temporali delle funzioni di fase, prese su un intervallo di tempo sufficientemente lungo (al limite infinito), corrisponderanno ai loro valori nello stato di equilibrio. Se fosse possibile dimostrare che le medie temporali delle funzioni di fase

$$\tilde{\varphi} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{+T} \varphi(p(t), q(t))\, dt$$

coincidono con le medie sulla superficie di energia Σ nello spazio delle fasi date dall'insieme microcanonico $\bar{\varphi} = \int_\Sigma \varphi(p, q)\, d\mu$, allora si sarebbe trovata una giustificazione di tipo fondamentale per l'impiego della teoria degli insiemi canonici nella meccanica statistica dell'equilibrio.

Boltzmann introdusse una classe di sistemi meccanici aventi la proprietà che l'orbita dei loro punti rappresentativi passa attraverso ogni punto della superficie di energia. Sistemi di tal genere sono chiamati sistemi ergodici (dal greco ἔργον 'lavoro', qui usato nel senso di energia, e ὁδός 'cammino'). Un sistema ergodico avrebbe le seguenti proprietà. Anzitutto tutte le orbite su una data superficie di energia dovrebbero essere identiche, poiché ogni orbita passa attraverso tutti i punti sulla superficie dell'energia e inoltre la direzione dell'orbita è determinata in ogni punto. Quindi le orbite effettive possono differire solo per l'istante in cui ogni punto è raggiunto. È immediato verificare che per i sistemi ergodici la media temporale $\tilde{\varphi}$ di ogni funzione di fase φ coincide con la media $\bar{\varphi}$ presa sulla superficie dell'energia secondo la misura invariante $d\mu$ dell'insieme microcanonico. Infatti, poiché la misura microcanonica è invariante, se il sistema è ergodico si ha $\varphi = \bar{\tilde{\varphi}}$. D'altra parte $\bar{\tilde{\varphi}} = \tilde{\bar{\varphi}}$, poiché le medie si possono liberamente scambiare. Ma è anche $\tilde{\bar{\varphi}} = \bar{\varphi}$, poiché l'insieme microcanonico è stazionario. Quindi in definitiva $\tilde{\varphi} = \bar{\varphi}$. In conclusione risulterebbe dimostrato che per i sistemi ergodici le medie sull'insieme microcanonico darebbero i valori corretti nella condizione di equilibrio.

Purtroppo fu ben presto compreso che sistemi strettamente ergodici nel senso descritto non possono esistere (una dimostrazione rigorosa fu data da Plancherel nel 1913), mentre l'introduzione di una ipotesi quasi-ergodica da parte degli Ehrenfest, secondo cui ogni orbita sulla superficie dell'energia passa arbitrariamente vicina ad ogni punto a patto di far trascorrere un tempo abbastanza lungo, non è sufficiente a dimostrare l'uguaglianza $\tilde{\varphi} = \bar{\varphi}$.

Le successive ricerche nell'ambito della teoria ergodica hanno teso alla precisazione delle condizioni di natura meccanica che il sistema deve soddisfare affinché si abbia l'uguaglianza tra le medie temporali e le medie in fase. Contributi importanti sono stati dati da Fermi, Birkhoff, Neumann, Hopf, Kolmogorov, Chinčin, Sinaj e molti altri [per gli sviluppi piú recenti, cfr. Arnol'd e Avez 1967; Lebowitz e Penrose 1973].

Tuttavia è chiaro che un'impostazione puramente basata sulla teoria ergodica non può dare una soluzione completamente soddisfacente al problema dei fondamenti della meccanica statistica. In effetti la teoria ergodica permette di trattare essenzialmente sistemi in condizioni di equilibrio, poiché essa si basa su medie temporali su periodi molto lunghi e su insiemi stazionari. Per lo studio dell'evoluzione irreversibile di sistemi fuori dall'equilibrio, occorre invece rivolgersi verso metodi piú adeguati, in grado di far raggiungere un duplice scopo: da un lato quello di rendere conto del carattere di irreversibilità dell'evoluzione e dall'altro quello di permettere uno studio quantitativo dei fenomeni irreversibili e della loro tendenza verso l'equilibrio, descritta dalle equazioni cinetiche.

Dal punto di vista storico il primo metodo, utilizzato con successo a questo proposito, fu quello della teoria cinetica dei gas, basato sul concetto di funzione di distribuzione delle velocità, che condusse alla celebre equazione di Boltzmann in grado di connettere le proprietà dell'evoluzione macroscopica di un gas e la sua struttura microscopica a livello molecolare.

Nel seguito si darà una descrizione di questi metodi, delle difficoltà incon-

trate e degli sviluppi successivi, con particolare riferimento al problema della compatibilità tra le proprietà dinamiche di un sistema meccanico e la sua evoluzione irreversibile dal punto di vista macroscopico.

5. *L'equazione di Boltzmann e i paradossi « Umkehreinwand » e « Wiederkehreinwand ».*

Prima d'introdurre la discussione sull'equazione cinetica di Boltzmann e sulle sue conseguenze, è opportuno illustrare brevemente il modo in cui si possono esprimere le grandezze fenomenologiche locali, introdotte nel § 2, in termini delle caratteristiche specifiche del livello di descrizione microscopico offerto dalla teoria cinetico-molecolare della materia.

In un mezzo continuo, non uniforme e variabile nel tempo, la densità $\rho(\mathbf{x}, t)$ in un generico punto $\mathbf{x}$ all'istante t può essere definita come il limite, per $\Delta\mathbf{x}$ tendente a zero, della densità media $\Delta m/\Delta\mathbf{x}$, dove Δm è la massa contenuta all'istante t in un piccolo volumetto $\Delta\mathbf{x}$ intorno al punto $\mathbf{x}$. Naturalmente questa definizione non può essere applicata nel caso in cui si prenda in considerazione un corpo materiale, per esempio un gas contenuto in un dato volume, di cui si è postulata una struttura discreta di tipo molecolare. Infatti in tal caso, prendendo il limite in senso matematico come descritto precedentemente, si avrebbe un valore della densità che varia rapidamente da punto a punto e da istante a istante e non corrisponde a nessuna grandezza fisica effettivamente misurabile. È questa una manifestazione della dicotomia continuo/discreto risultante dalla rappresentazione microscopica di tipo molecolare per un mezzo che è descritto da variabili continue a livello macroscopico. Il modo corretto di procedere costituisce un ottimo esempio di come vanno considerati i processi di limite nell'ambito di considerazioni di tipo fisico. Basta infatti scegliere un volumetto $\Delta\mathbf{x}$ intorno al punto $\mathbf{x}$ che sia abbastanza grande in modo da contenere un grandissimo numero di molecole, ma abbastanza piccolo rispetto alle scale tipiche di variabilità delle grandezze macroscopiche che descrivono localmente lo stato del sistema termodinamico (densità, pressione, temperatura, densità di energia interna, ecc.). Per rendersi conto che è possibile, nei casi piú generali, realizzare questo compromesso, basta considerare che un cubetto con lo spigolo di un centesimo di millimetro contiene circa ventisette miliardi di molecole di un gas a pressione atmosferica e temperatura ambiente. Si consideri la massa Δm di gas contenuta mediamente nel volumetto $\Delta\mathbf{x}$ durante un intervallo di tempo Δt, che sia grande rispetto al tempo medio impiegato da una molecola a percorrere se libera una distanza uguale alle dimensioni lineari del volumetto $\Delta\mathbf{x}$, ma contemporaneamente sia piccolo rispetto alla scala di variazione temporale delle proprietà macroscopiche del mezzo. Anche questo compromesso è realizzabile in generale. Per esempio, poiché la velocità media delle molecole in un gas nelle condizioni standard prima specificate è di parecchie centinaia di metri al secondo, una molecola se libera percorrerebbe la distanza di un centesimo di millimetro in meno di un decimo di milionesimo di secondo, che è

un intervallo di tempo estremamente piú piccolo della scala tipica con cui è possibile seguire istante per istante l'evoluzione delle proprietà macroscopiche locali del gas. Cosí specificato $\Delta\mathbf{x}$ e Δm, si assumerà semplicemente per la densità $\rho(\mathbf{x}, t)$ il valore del rapporto $\Delta m/\Delta\mathbf{x}$, *senza* effettuare ulteriori processi di limite.

Piú in generale si può osservare che le molecole contenute nel volumetto $\Delta\mathbf{x}$ saranno distribuite secondo varie possibili velocità, per cui è conveniente introdurre una funzione di distribuzione $f(\mathbf{x}, \mathbf{v}; t)$ il cui significato fisico è che $f(\mathbf{x}, \mathbf{v}; t)\,\Delta\mathbf{x}\,\Delta\mathbf{v}$ rappresenta il numero medio di molecole contenute nel volumetto $\Delta\mathbf{x}$ ed aventi velocità contenute in un elementino di volume $\Delta\mathbf{v}$, nello spazio delle velocità, dove le medie s'intendono fatte nel corso di un intervallo di tempo Δt intorno all'istante t. Analogamente si possono introdurre funzioni di distribuzione che tengono conto delle correlazioni multiple, cosí per esempio $f^{(s)}(\mathbf{x}_1, \ldots, \mathbf{x}_s, \mathbf{v}_1, \ldots, \mathbf{v}_s; t)\,\Delta\mathbf{x}_1 \ldots \Delta\mathbf{x}_s\,\Delta\mathbf{v}_1 \ldots \Delta\mathbf{v}_s$ è connessa con la probabilità di trovare s molecole negli elementi di volume $\Delta\mathbf{x}_1, \ldots, \Delta\mathbf{x}_s$ con velocità in $\Delta\mathbf{v}_1, \ldots, \Delta\mathbf{v}_s$, rispettivamente, all'istante t. In un gas a densità ordinaria, con sufficiente approssimazione,

$$f^{(s)}(\mathbf{x}_1, \ldots, \mathbf{x}_s, \mathbf{v}_1, \ldots, \mathbf{v}_s; t) = \prod_{i=1}^{s} f(\mathbf{x}_i, \mathbf{v}_i; t),$$

ma in un gas denso questa approssimazione non è piú adeguata, poiché ogni molecola influisce direttamente, tramite il suo campo di forza, sulla distribuzione delle altre molecole in sua prossimità.

È evidente, sulla base dell'ipotesi cinetico-molecolare, che la funzione di distribuzione f permette di definire, almeno nel caso di gas a bassa densità, tutti i parametri fenomenologici macroscopici locali che individuano istante per istante lo stato del sistema termodinamico (cfr. § 2). Per esempio la densità sarà data da

$$\rho(\mathbf{x}, t) = m \int f(\mathbf{x}, \mathbf{v}; t)\, d\mathbf{v},$$

dove m è la massa di ogni singola molecola del gas e l'integrale s'intende esteso a tutte le possibili velocità. Analogamente, se per una qualsiasi grandezza $A(\mathbf{x}, \mathbf{v}; t)$ s'introduce la media parziale sulle velocità definita da

$$\langle A \rangle(\mathbf{x}, t) = \frac{\int A(\mathbf{x}, \mathbf{v}; t)\, f(\mathbf{x}, \mathbf{v}; t)\, d\mathbf{v}}{\int f(\mathbf{x}, \mathbf{v}; t)\, d\mathbf{v}}$$

si possono esprimere gli altri parametri locali, introdotti nel § 2, nella forma

$$u_i(\mathbf{x}, t) = \langle v_i \rangle$$

per il campo di velocità media di flusso,

$$\varepsilon(\mathbf{x}, t) = \frac{1}{2}\rho(\mathbf{x}, t)\langle \mathbf{w}^2 \rangle$$

per la densità di energia interna,

$$P_{ij}(\mathbf{x}, t) = \rho(\mathbf{x}, t)\langle w_i w_j \rangle,$$

per il tensore degli sforzi,

$$q_i(\mathbf{x}, t) = \frac{1}{2}\rho(\mathbf{x}, t)\langle w_i \mathbf{w}^2 \rangle$$

per la densità del flusso di corrente del calore, dove $\mathbf{w}(\mathbf{x}, \mathbf{v}; t) = \mathbf{v} - \mathbf{u}(\mathbf{x}, t)$ è la cosiddetta velocità termica.

Naturalmente nel caso di gas ad alta densità i parametri fenomenologici locali coinvolgeranno anche le funzioni di distribuzione superiori $f^{(2)}$, ecc., ma per gli scopi del presente articolo è sufficiente la situazione delineata precedentemente, valida nel caso di gas a bassa densità.

La discussione precedente mostra chiaramente come nel passaggio dalla descrizione a livello microscopico, in cui lo stato del sistema meccanico è determinato da posizioni e velocità di tutte le molecole, alla descrizione a livello macroscopico, in cui lo stato del sistema è fenomenologicamente determinato dalla funzione di distribuzione f lentamente variabile nello spazio e nel tempo, si sia avuta una sorta di contrazione che ha sostanzialmente ridotto il numero dei gradi di libertà, tramite l'introduzione di considerazioni di natura statistica. È altresí chiaro che uno stesso stato macroscopico può essere ottenuto da piú stati microscopici, poiché i dettagli piú fini della descrizione microscopica vengono persi nel corso del procedimento di medie spaziali e temporali. Si vedrà che questa contrazione di variabili, componente essenziale del metodo statistico, svolge un grande ruolo nell'instaurarsi di proprietà di irreversibilità nella descrizione macroscopica.

Poiché i parametri fenomenologici locali dipendono dalla funzione di distribuzione f, è chiaro che per ottenere le equazioni di evoluzione macroscopiche è necessario conoscere come f varia nel tempo. Attraverso una lucida e coraggiosa analisi, partendo da ipotesi estremamente ragionevoli, Boltzmann derivò un'equazione di evoluzione di tipo integrodifferenziale per la funzione di distribuzione f, come conseguenza dell'evoluzione meccanica del sottostante modello molecolare.

A causa delle notevoli complicazioni tecniche, non è purtroppo possibile fornire, nei limiti di questa esposizione, una derivazione completa di questa equazione, per la quale si rimanda alle lezioni di Boltzmann [1896-98] o alla piú recente esposizione di Grad [1958]. Verrà però descritto nel prossimo paragrafo un modello, dovuto agli Ehrenfest [1912], che contiene in forma elementare tutte le caratteristiche e i paradossi dell'equazione di Boltzmann.

Nel caso in cui le molecole del gas sono assimilate a piccole sferette elastiche, allora la forma caratteristica dell'equazione di Boltzmann è

$$\frac{\partial f}{\partial t} + (\mathbf{v}\cdot\nabla)f + (\mathbf{F}\cdot\nabla_{\mathbf{v}})f + \int (ff_1 - f'f_1')\, a\, d\mathbf{v}_1\, d\boldsymbol{\omega} = 0$$

dove, oltre ai simboli dal significato evidente, $\mathbf{F}$ è la forza esterna agente sulla singola molecola nella posizione $\mathbf{x}$, $\nabla_{\mathbf{v}}$ è il gradiente ottenuto prendendo le derivate rispetto alle componenti della velocità, $\mathbf{v}$, $\mathbf{v}_1$ sono le velocità di due sfere prima dell'urto e $\mathbf{v}'$, $\mathbf{v}_1'$ quelle dopo, mentre $\boldsymbol{\omega}$ è il versore individuato dai centri delle sferette in collisione e a è una grandezza cinematica dell'urto. Nell'equazione sono state utilizzate le seguenti abbreviazioni: $f = f(\mathbf{x}, \mathbf{v}; t)$, $f_1 = f(\mathbf{x}, \mathbf{v}_1; t)$, $f' = f(\mathbf{x}, \mathbf{v}'; t)$ e $f_1' = f(\mathbf{x}, \mathbf{v}_1'; t)$.

Nella sua struttura essenziale l'equazione cinetica descrive la variazione nel

tempo della funzione di distribuzione $f(\mathbf{x}, \mathbf{v}; t)$, dovuta al fatto che molecole abbandonano o raggiungono il volumetto $\Delta\mathbf{x}\Delta\mathbf{v}$ per effetto della loro velocità $\mathbf{v}$ e della loro accelerazione dovuta alla forza esterna $\mathbf{F}$ (di questo bilancio tengono conto il secondo e terzo termine) oppure per effetto degli urti con le altre molecole, di cui tiene conto l'ultimo termine.

La caratteristica piú importante dell'equazione di Boltzmann è che essa rende conto, almeno in linea di principio, di tutte le proprietà termodinamiche di tipo fenomenologico descritte nel precedente § 2, dalla proprietà di esistenza di uno stato di equilibrio termodinamico, alla legge di accrescimento dell'entropia, alle leggi cinetiche dei fenomeni di non-equilibrio, come mostrarono lo stesso Boltzmann e i suoi continuatori.

Dati i limiti della presente trattazione, non è possibile dare una derivazione approfondita di queste rilevanti conseguenze dell'equazione di Boltzmann. Ci si limiterà pertanto ad una rapida e schematica illustrazione dei punti piú importanti per il loro aspetto fisico.

Si osservi anzitutto che l'equazione di Boltzmann è del primo ordine rispetto al tempo e pertanto si pone per essa un problema di valori iniziali, cioè quello di trovare la funzione di distribuzione in ogni istante $t \geq 0$ dato il suo valore nell'istante iniziale $f(\mathbf{x}, \mathbf{v}; 0)$. Fu un grande trionfo per Boltzmann l'essere in grado di dimostrare, non con pieno rigore matematico, ma con convincente vigore fisico, che ogni distribuzione iniziale evolve in modo da tendere nel corso del tempo verso la distribuzione canonica di equilibrio termodinamico, di cui si è trattato nel § 4. Pertanto l'approccio verso l'equilibrio è contenuto nella stessa equazione.

Un'altra importante conseguenza dell'equazione di Boltzmann è il famoso teorema H. Occorre introdurre la funzione del tempo

$$H(t) = \iint f(\mathbf{x}, \mathbf{v}; t) \log f(\mathbf{x}, \mathbf{v}; t)\, d\mathbf{x}\, d\mathbf{v}$$

nota come funzione H di Boltzmann.

Mediante semplici manipolazioni si può subito dimostrare che l'equazione di Boltzmann implica che la funzione H è sempre decrescente nel tempo, a meno che il sistema non si trovi in uno stato di equilibrio termodinamico descritto da una funzione di distribuzione di tipo canonico, in questo caso la funzione H resta costante nel tempo (teorema H).

Da queste proprietà della funzione H Boltzmann dedusse la sua stretta connessione con l'entropia termodinamica tramite la relazione d'identificazione $S = -KH$, dove K è la costante di Boltzmann. In questo modo l'affermazione del teorema H secondo cui, per effetto delle collisioni, la funzione H non può mai crescere nel tempo, viene a trovarsi in corrispondenza stretta con la ben nota affermazione fenomenologica del secondo principio della termodinamica in base al quale l'entropia non può mai decrescere.

Si vede cosí che l'equazione di Boltzmann, almeno entro i suoi limiti di validità, fornisce un'interpretazione completa su basi cinetico-molecolari alle piú importanti affermazioni della termodinamica fenomenologica.

Tra le applicazioni piú importanti dal punto di vista storico dell'equazione di Boltzmann possono essere ricordate le derivazioni delle espressioni teoriche per i coefficienti di viscosità, la conducibilità termica e altre proprietà cinetiche dei gas, ottenute nel 1916-17 da Chapman e Enskog, che lavorarono indipendentemente utilizzando potenti metodi di approssimazione, conseguenti ad una prima analisi fatta da Hilbert, che era stato significativamente attratto dall'estremo interesse sollevato dai problemi matematici connessi con l'equazione di Boltzmann.

Agli effetti della presente trattazione è invece importante porre in rilievo una caratteristica fondamentale dell'equazione di Boltzmann, quella di non possedere invarianza per inversione temporale. L'evoluzione temporale descritta è quindi tipicamente irreversibile e per questo può dar luogo alla spiegazione della legge di accrescimento dell'entropia e delle equazioni cinetiche che sono anch'esse non invarianti per inversione temporale.

Naturalmente, poiché, in ultima istanza, l'equazione di Boltzmann è «derivata» dalle leggi della meccanica, l'irreversibilità non può non produrre delle difficoltà nell'accettazione della derivazione. In particolare Loschmidt propose il suo famoso *Umkehreinwand*, il paradosso della reversibilità, che nell'ambito del quadro tracciato precedentemente può essere descritto come segue. Si supponga di considerare il sistema meccanico invertito temporalmente, con $t \to t' = -t$, $\mathbf{x}_s \to \mathbf{x}'_s(t') = \mathbf{x}_s(t)$, $\mathbf{v}_s \to \mathbf{v}'_s(t) = -\mathbf{v}_s(t)$, dove $\mathbf{x}_s$ e $\mathbf{v}_s$ sono la posizione e la velocità della s-esima molecola. A causa dell'invarianza delle equazioni del moto descritta precedentemente, è chiaro che $\mathbf{x}'_s$, $\mathbf{v}'_s$ percorreranno le loro traiettorie in funzione di t' con pieno rispetto delle leggi meccaniche del moto. D'altra parte, per come è stata definita la funzione di distribuzione f in termini delle grandezze microscopiche, essa deve trasformarsi nel modo seguente: $f \to f'(\mathbf{x}', \mathbf{v}'; t') = f(\mathbf{x}, -\mathbf{v}; t)$. Tuttavia è possibile immediatamente verificare che f' *non* soddisfa l'equazione di Boltzmann, ma soddisfa un'equazione simile con il segno cambiato per l'ultimo termine, con conseguenze disastrose dal punto di vista fisico. Infatti la nuova equazione prevederebbe un abbandono spontaneo dello stato di equilibrio termodinamico sotto piccole perturbazioni, la violazione del secondo principio della termodinamica, la conseguenza che l'attrito dovrebbe agevolare il flusso dei fluidi invece di ostacolarlo, ecc.

Dal punto di vista del teorema H il paradosso di reversibilità può essere formulato nel modo seguente. Si consideri un sistema meccanico il cui stato a ciascun istante t sia rappresentato da un punto $P(t)$ nello spazio delle fasi Γ. In corrispondenza di una serie successiva di istanti, ..., t_1, t_2, ..., t_n, ..., con $t_i < t_j$ per $i < j$, il punto rappresentativo attraversa le sequenza di punti nello spazio Γ ..., P_1, P_2, ..., P_n, ...

Secondo il teorema H si avrà allora la seguente sequenza per i corrispondenti valori della funzione H: $\ldots \geq H_1 \geq H_2 \geq \ldots \geq H_n \geq \ldots$ (il segno di uguaglianza vale solo se la funzione di distribuzione è canonica).

Si consideri ora la stessa traiettoria percorsa in senso inverso. Poiché sulla base della sua definizione la funzione H non cambia per inversione temporale, si ottiene immediatamente che nel corso del moto invertito la funzione H passa

successivamente attraverso i valori $\ldots \leq H_n \leq \ldots \leq H_2 \leq H_1 \leq \ldots$ Ci si rende allora conto del fatto che per ogni moto del sistema per il quale la funzione H decresce uniformemente da H_1 ad H_n esiste un altro moto dello stesso sistema per il quale la funzione H cresce esattamente allo stesso modo dal valore H_n al valore H_1. Data la connessione della funzione H con l'entropia si avrebbe una violazione del secondo principio della termodinamica.

Si vede quindi che la reversibilità temporale microscopica pone seriamente in crisi il teorema H e quindi l'intera riduzione, operata da Boltzmann, della termodinamica fenomenologica a conseguenza delle proprietà dinamiche delle molecole costituenti in ultima istanza il sistema fisico considerato.

Un'altra seria obiezione fu successivamente sollevata da Zermelo con il suo paradosso di ricorrenza (*Wiederkehreinwand*), basato su un teorema di Poincaré relativo a sistemi contenuti in un volume finito. Poincaré aveva dimostrato che se un sistema meccanico passa attraverso la sequenza di punti P_1, P_2, ..., P_n negli istanti t_1, t_2, ..., t_n, allora questa sequenza verrà ripetuta, approssimativamente ma con accuratezza comunque grande, nel corso dell'evoluzione naturale ulteriore, a patto di aspettare un tempo abbastanza lungo. In altre parole, specificato un margine di tolleranza sui valori delle posizioni e degli impulsi comunque piccolo, esiste un intervallo di tempo T tale che il sistema, dopo aver percorso P_1, P_2, ..., P_n negli istanti t_1, t_2, ..., t_n, si troverà a percorrere una sequenza P'_1, P'_2, ..., P'_n negli istanti $T+t_1$, $T+t_2$, ..., $T+t_n$, dove i punti P'_i scartano rispettivamente dai punti P_i nei limiti della tolleranza scelta. In genere volendo una precisione maggiore occorrerà attendere un tempo T maggiore. I valori della funzione H corrispondenti ai punti P'_i saranno H'_1, H'_2, ..., H'_n, approssimativamente uguali, nei limiti della tolleranza consentita che può essere molto piccola, ai valori H_1, ..., H_n. Se quindi, sulla base del teorema H, si ha $H_1 > H_n$, si ottiene anche $H'_1 > H_n$, in contraddizione con il teorema H poiché H_n corrisponde all'istante t_n e H'_1 all'istante successivo $T+t_1$.

È però importante notare che il tempo di ricorrenza nel caso di concreti sistemi termodinamici può essere estremamente lungo. Per esempio Boltzmann prese in considerazione un centimetro cubo di gas in condizioni normali di pressione e temperatura e valutò che il tempo che occorrerebbe attendere in media perché tutti gli atomi riprendano le loro posizioni iniziali, a meno di errori dell'ordine di un milionesimo di millimetro, e le loro velocità iniziali, a meno di errori dell'ordine di un metro al secondo, è dell'ordine di $10^{10^{19}}$ anni!

Le obiezioni di Loschmidt (*Umkehreinwand*) e di Zermelo (*Wiederkehreinwand*) erano estremamente ben fondate dal punto di vista matematico, per cui minacciavano di far crollare completamente l'edificio concettuale costruito da Boltzmann, il cui contenuto principale era la derivazione delle leggi della termodinamica fenomenologica dalle proprietà meccaniche dei modelli della teoria cinetico-molecolare, con conseguenze molto gravi per l'intera ipotesi molecolare.

Pertanto Boltzmann ed i suoi seguaci furono indotti a perseguire una dettagliata analisi della derivazione dell'equazione cinetica dalle equazioni della meccanica, che ebbe come risultato quello di isolare gli aspetti di natura sta-

tistica, implicitamente introdotti nella derivazione, che in particolare permettevano il passaggio da una situazione di completa reversibilità ad una di essenziale irreversibilità temporale.

Venne quindi riconosciuto che il ruolo centrale nella derivazione era svolto da un'assunzione fondamentale sul numero probabile di collisioni (*Stosszahlansatz*) che le molecole con determinate velocità e posizioni subiscono per effetto della presenza delle altre molecole. Questa assunzione non poteva certamente essere giustificata da considerazioni di natura puramente meccanica, ma aveva il carattere di postulato di natura statistica.

Di conseguenza Boltzmann, con una intuizione veramente geniale, riuscí a capovolgere la forza dell'argomento, riconoscendo che la validità dell'equazione da lui derivata, e delle sue conseguenze, tra cui il teorema H, non aveva un carattere assoluto, ma solo un carattere di elevatissima probabilità, nel senso che tutte le situazioni microscopiche che conducono alla corretta evoluzione macroscopica hanno una piú elevata probabilità di verificarsi rispetto alle situazioni microscopiche di tipo patologico che comportano innaturali evoluzioni macroscopiche.

In questo modo possono essere evitate le conseguenze negative dei paradossi di reversibilità e di ricorrenza. Infatti questi paradossi si basano su teoremi della meccanica che si riferiscono agli stati dinamici esatti del sistema a livello microscopico. Invece le considerazioni basate sull'equazione di Boltzmann e sul teorema H coinvolgono una descrizione macroscopica del sistema, basata sulla funzione di distribuzione $f(\mathbf{x}, \mathbf{v}; t)$, alla quale corrisponde un numero molto grande di stati microscopici, alcuni dei quali possono comportare anche un innaturale incremento della funzione H. Le valutazioni di natura statistica sul numero piú probabile di collisioni (*Stosszahlansatz*) mostrano che il numero di queste ultime è molto piú piccolo in relazione agli stati microscopici che corrispondono a un decremento della funzione H, almeno se il sistema è abbastanza lontano dall'equilibrio. Come conseguenza, il decremento della funzione H è molto piú probabile che non un incremento.

Il paradosso di reversibilità invece è basato sulla considerazione di un singolo stato microscopico del sistema, per cui l'effettiva realizzazione dell'evoluzione invertita temporalmente sarebbe possibile solo a condizione che le velocità di tutte le molecole fossero accuratamente note e potessero essere invertite ad un dato istante. Se s'introduce il concetto di preparazione del sistema in un dato stato macroscopico, cioè la costruzione effettiva di stati microscopici che corrispondono ad una determinata situazione macroscopica, allora il teorema H di Boltzmann esprime semplicemente il fatto che è molto piú semplice, quando lo stato macroscopico considerato è sufficientemente lontano dall'equilibrio, preparare il sistema in uno stato che dà origine a un decremento di H piuttosto che riuscire a realizzare condizioni iniziali per le quali H cresce.

Il riconoscimento del carattere statistico dell'equazione di Boltzmann evita quindi ogni contraddizione con il carattere reversibile dell'evoluzione meccanica a livello microscopico.

Neanche le affermazioni contenute nel paradosso di ricorrenza sono in contraddizione con il fatto che l'evoluzione piú probabile del sistema a un determinato istante comporta un decremento della funzione H, poiché la natura statistica della validità del teorema H non esclude la possibilità di avere incrementi della funzione H per alcuni stati. Al contrario le considerazioni di natura statistica e di natura meccanica si complementano reciprocamente e permettono di descrivere l'evoluzione della funzione H come una serie di fluttuazioni intorno al valore minimo [Ehrenfest 1912]. Inoltre queste fluttuazioni sono meno probabili quanto piú esse deviano dalla posizione di equilibrio, tanto da risultare in questi casi addirittura impossibili ad osservarsi, anche a causa della durata enorme del tempo di ricorrenza in tutte le situazioni concrete.

Il punto di vista di Boltzmann, essenzialmente corretto, ma la cui natura non è stata ancora completamente chiarita, permise un'interpretazione completamente nuova, di natura probabilistica, del secondo principio della termodinamica e in generale dei fenomeni di irreversibilità.

La reversibilità microscopica porta quindi ad un'altamente probabile irreversibilità macroscopica. Questa linea di soluzione della dicotomia reversibilità/irreversibilità verrà ulteriormente analizzata in seguito.

Per concludere questo paragrafo si ricorda che contributi molto importanti per una completa comprensione della derivazione dell'equazione di Boltzmann, o piú in generale delle equazioni cinetiche fenomenologiche, dai principî della meccanica sono stati dati da Yvon [1935], Bogoljubov [1946], Born e Green [1949], Kirkwood [1946; 1947] e Prigogine [1962], oltre che da moltissimi altri ricercatori, attraverso una serie notevole di investigazioni che hanno sostanzialmente confermato la correttezza delle intuizioni originarie di Boltzmann.

6. *Il modello degli Ehrenfest.*

Allo scopo di esaminare in dettaglio alcuni degli aspetti introdotti nella precedente sezione, è conveniente discutere le caratteristiche di un modello semplificato, introdotto dagli Ehrenfest [1912], che non richiede l'uso di complicati strumenti formali.

Si consideri un grande numero N di punti materiali, le molecole P, che si muovono liberamente, senza interagire tra di loro, su un piano infinito, riferito agli assi coordinati (x, y). S'introducano poi le molecole Q, costituite da quadratini fissi di spigolo a e distribuite a caso in maniera uniforme sull'intero piano in modo che la distanza media tra di esse, A, sia molto grande rispetto alle dimensioni a. Si suppone che ogni molecola Q sia disposta in modo che le sue diagonali siano esattamente parallele agli assi x e y e che inoltre le molecole P, incontrando le molecole Q, subiscano un urto elastico.

Si assume che ad un certo istante tutte le molecole P abbiano la stessa velocità c e che si muovano lungo una delle seguenti possibili quattro direzioni: l'asse x positivo $\rightarrow$ (1), l'asse y positivo $\uparrow$ (2), l'asse x negativo $\leftarrow$ (3) e l'asse y negativo $\downarrow$ (4) (cfr. fig. 1). A causa della natura del sistema e delle interazioni permesse, è chiaro che la situazione in ogni altro istante sarà perfettamente si-

mile, dal punto di vista qualitativo, a quella descritta. L'unica differenza possibile, a parte la diversa posizione sul piano delle molecole P, sarà costituita dal numero di molecole P che si muovono nelle quattro direzioni permesse.

Si ha qui la possibilità d'introdurre in maniera estremamente chiara i due livelli di descrizione, microscopico e macroscopico, che sono stati trattati nelle pagine precedenti introducendo le idee fondamentali della teoria cinetico-molecolare della materia.

A livello microscopico lo stato del sistema in un determinato istante è completamente descritto dando la posizione di ogni singola molecola P e la direzione lungo la quale essa si muove (si ricordi che la velocità è assunta costante). Per seguire l'evoluzione del sistema negli istanti successivi basta considerare il moto di ogni singola molecola P. Questo sarà rettilineo uniforme nell'intervallo di tempo tra due urti successivi, quando la molecola P incontra una molecola Q essa cambierà bruscamente di direzione passando per esempio dalla direzione 1 alla 2 o alla 4, ecc. Naturalmente la descrizione microscopica è completamente reversibile nel tempo. Se ad un certo istante s'inverte la direzione di una molecola P, essa ripercorrerà fedelmente in senso inverso la traiettoria già percorsa.

Per la descrizione macroscopica del sistema si possono assumere i quattro numeri f_1, f_2, f_3, f_4, i quali rappresentano in ogni istante t il numero medio di molecole che si muovono in ciascuna delle quattro direzioni permesse, dove la media s'intende presa su un intervallino Δt, intorno all'istante t, abbastanza grande rispetto all'intervallo di tempo medio tra gli urti T_1 (cioè durante Δt deve avvenire un numero di urti molto grande) ma abbastanza piccolo rispetto al tempo caratteristico T_2 di variabilità macroscopica delle funzioni di distribuzione f_i. Ovviamente risulta $f_1+f_2+f_3+f_4=N$. Si noti che di tutte le informazioni contenute nella descrizione microscopica solo una piccola parte di esse

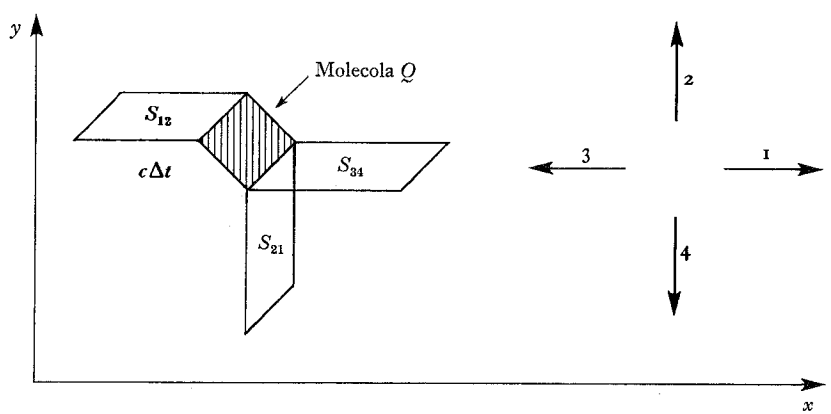


Figura 1.
Effetto degli urti delle molecole P contro molecole Q.

è andata a confluire nella descrizione macroscopica, e che inoltre a una stessa situazione macroscopica possono corrispondere moltissime situazioni microscopiche. Questo chiarisce il carattere «contratto» della descrizione macroscopica. Si noti anche l'importanza della media temporale che introduce una sorta di raddolcimento nella dipendenza temporale delle f_i e permette di scrivere equazioni differenziali per la loro evoluzione.

Le funzioni di distribuzione $f_i(t)$ $(i=1, 2, 3, 4)$ svolgono lo stesso ruolo della funzione di distribuzione $f(\mathbf{x}, \mathbf{v}; t)$ della teoria cinetica dei gas. Ovviamente, per ragioni di simmetria, la distribuzione di equilibrio, corrispondente alla distribuzione di Maxwell, è data da $f_1^0=f_2^0=f_3^0=f_4^0=N/4$. Occorre dimostrare che, sotto determinate assunzioni probabilistiche, corrispondenti allo *Stosszahlansatz* per l'equazione di Boltzmann, ogni distribuzione f_i, diversa da f_i^0, tenderà nel tempo verso la distribuzione di equilibrio f_i^0, come conseguenza degli urti delle molecole P con molecole Q. Inoltre la distribuzione di equilibrio dovrà mantenersi costante nel tempo.

Si indichi con $N_{12}\Delta t$ il numero di molecole P che per effetto degli urti passano dalla direzione 1 alla direzione 2 nell'intervallo di tempo Δt. Queste sono tutte e solo quelle che all'inizio dell'intervallo temporale Δt soddisfano simultaneamente le seguenti due condizioni: *a*) si muovono nella direzione 1, *b*) si trovano in una delle strisce S_{12}, costituite da parallelogrammi di lunghezza $c\Delta t$ con una delle basi appoggiate su ciascuna delle molecole Q dalla parte del lato che per effetto dell'urto fa deviare le molecole P provenienti dalla direzione 1 nella direzione 2 (cfr. fig. 1).

È chiaro che la conoscenza dei numeri f_1, f_2, f_3, f_4 non è sufficiente a determinare quante molecole P soddisfano le condizioni A e B. Pertanto se si vuol descrivere l'evoluzione macroscopica tramite le sole variabili f_i bisogna ricorrere ad assunzioni di tipo probabilistico che permettono di determinare mediamente il numero $N_{12}\Delta t$, senza introdurre ulteriori informazioni provenienti dalla piú dettagliata descrizione microscopica introdotta precedentemente.

In analogia con il caso di Boltzmann, si introduce la seguente assunzione sul numero di urti (*Stosszahlansatz*): La frazione di molecole P di ogni singola direzione che si vengono a trovare mediamente nelle strisce S_{ij} è la stessa che il rapporto tra l'area totale delle strisce e l'area totale libera del piano. Si indica questo rapporto con $k\Delta t$, dove $k=nca/\sqrt{2}$ e n è la densità di molecole Q sul piano (notare che l'area di ogni striscia S_{ij} è $c\Delta t a/\sqrt{2}$ e che il numero medio di strisce S_{ij} presenti su una grande superficie Σ è dato da $n\Sigma$).

Si ha quindi in particolare $N_{12}\Delta t=kf_1\Delta t$, e analogamente per le altre transizioni. Pertanto il numero medio di collisioni nell'intervallo Δt è dato da

$$(N_{12}+N_{21}+N_{23}+N_{32}+N_{34}+N_{43}+N_{41}+N_{14})\Delta t=$$
$$=k(f_1+f_2+f_2+f_3+f_3+f_4+f_4+f_1)\Delta t=2Nk\Delta t,$$

e quindi l'intervallo medio tra due collisioni qualsiasi è

$$T_1=\frac{1}{2Nk}=\frac{1}{\sqrt{2}acnN}.$$

Sulla base delle assunzioni poste dallo *Stosszahlansatz* si possono immediatamente ricavare le equazioni di evoluzione macroscopiche per le f_i. Tenendo conto del fatto che per esempio f_1 subisce nell'intervallo di tempo Δt una variazione Δf_1 dovuta sia alle perdite per gli urti $1 \to 2$ e $1 \to 4$ sia ai guadagni per gli urti $2 \to 1$ e $4 \to 1$, si ha

$$\Delta f_1 = (-N_{12} - N_{14} + N_{21} + N_{41})\Delta t = k(f_2 + f_4 - 2f_1)\Delta t,$$

e analogamente per le altre f_i. Considerando intervalli Δt abbastanza piccoli, è possibile scrivere immediatamente le seguenti equazioni differenziali per le f_i:

$$\frac{df_1}{dt} = k(f_2 + f_4 - 2f_1)$$

$$\frac{df_2}{dt} = k(f_3 + f_1 - 2f_2)$$

$$\frac{df_3}{dt} = k(f_4 + f_2 - 2f_3)$$

$$\frac{df_4}{dt} = k(f_1 + f_3 - 2f_4),$$

Esse sono l'analogo dell'equazione di Boltzmann nel caso di un gas. Da queste equazioni si vede immediatamente che la distribuzione di equilibrio $f_i^0 = N/4$ è stazionaria, infatti in tal caso i secondi membri si annullano e non ci sono variazioni nel tempo. Inoltre è facile mostrare che ogni distribuzione iniziale tende verso la distribuzione di equilibrio. Basta combinare opportunamente le equazioni, sottraendo la terza dalla prima oppure la quarta dalla seconda oppure sommando la prima e la terza e sottraendo la seconda e la quarta, per ottenere

$$\frac{d(f_1 - f_3)}{dt} = -2k(f_1 - f_3),$$

$$\frac{d(f_2 - f_4)}{dt} = -2k(f_2 - f_4),$$

$$\frac{d[(f_1 + f_3) - (f_2 + f_4)]}{dt} = -4k[(f_1 + f_3) - (f_2 + f_4)].$$

Poiché in generale la soluzione $y(t)$ dell'equazione $dy/dt = -\alpha y$ è $y(t) = y(0)e^{-\alpha t}$ e quindi se $\alpha > 0$, si ha che $y(t)$ tende ad annullarsi esponenzialmente per $t \to \infty$, si ha immediatamente che $f_1 - f_3 \to 0$, $f_2 - f_4 \to 0$ e $(f_1 + f_3) - (f_2 + f_4) \to 0$ per $t \to \infty$, cioè si ottiene una graduale tendenza di tipo esponenziale al livellamento delle quattro f_i verso la condizione di equilibrio.

Si è quindi visto come, in questo modello semplificato, ragionevoli assunzioni statistiche sulla distribuzione degli urti conducano a una descrizione macroscopica in accordo con i principî della termodinamica, perché è in grado di spiegare in particolare l'esistenza di uno stato macroscopico di equilibrio termodinamico e la naturale tendenza dei sistemi fuori dall'equilibrio a subire processi di natura irreversibile che li riconducono all'equilibrio.

Si noti che la scala macroscopica di variabilità temporale suggerita dalle equazioni (tempo di rilassamento) è dell'ordine di grandezza di $T_2 \simeq 1/k = \sqrt{2}/nca$. Risulta pertanto T_1 molto piú piccolo di T_2 e le assunzioni sulle medie temporali nella definizione delle funzioni di distribuzione f_i risultano pienamente giustificate a posteriori.

Per illustrare la validità del teorema H in questo caso particolare si può anzitutto introdurre la funzione H di Boltzmann definita da

$$H(t) = f_1 \log f_1 + f_2 \log f_2 + f_3 \log f_3 + f_4 \log f_4.$$

Si vede immediatamente che per effetto delle equazioni cinetiche per le f_i si ha

$$\frac{dH}{dt} = -k\left[(f_1 - f_2)\log\frac{f_1}{f_2} + (f_2 - f_3)\log\frac{f_2}{f_3} + (f_3 - f_4)\log\frac{f_3}{f_4} + (f_4 - f_1)\log\frac{f_4}{f_1}\right].$$

Poiché $(a-b)\log(a/b)$ è sempre positivo tranne nel caso in cui sia $a = b$, si ottiene immediatamente $dH/dt \leq 0$, dove il segno di uguaglianza vale solo se $f_1 = f_2 = f_3 = f_4$. Si ottiene quindi che la funzione H decresce nel tempo a meno che la distribuzione delle f non sia quella di equilibrio, in questo caso H resta stazionaria. Resta cosí dimostrata per via del tutto elementare la validità del teorema H nel caso del modello degli Ehrenfest.

Dal punto di vista della presente trattazione è importante porre in rilievo la natura essenzialmente irreversibile delle equazioni cinetiche per le f_i. Anzitutto si osservi che, per come sono state definite le funzioni di distribuzione, se si effettua un'inversione temporale, $t \to t' = -t$, allora si deve avere $f_1 \to f_1'(t') = f_3(t)$, $f_2 \to f_2'(t') = f_4(t)$, $f_3 \to f_3'(t') = f_1(t)$, $f_4 \to f_4'(t') = f_2(t)$, poiché tutte le velocità devono essere invertite. Ma allora si può immediatamente verificare che le f_i' soddisfano in funzione di t' equazioni cinetiche completamente diverse da quelle per le f, che si ottengono tramite la sostituzione $k \to -k$. Queste nuove equazioni cinetiche non mostrano comportamento termodinamico, in particolare una piccola fluttuazione fuori dallo stato di equilibrio viene immediatamente amplificata dalle equazioni e conduce il sistema verso stati macroscopici sempre piú lontani dall'equilibrio.

Si è visto cosí che le equazioni cinetiche precedentemente ottenute a partire dallo *Stosszahlansatz*, e il conseguente teorema H, non possono avere una validità illimitata, ma solo di tipo probabilistico, poiché in generale esistono stati microscopici che conducono a descrizioni macroscopiche con comportamento termodinamico inaccettabile. Sarebbe interessante caratterizzare piú in dettaglio questi stati microscopici patologici, facendo vedere che essi hanno, in un senso da specificarsi, scarsa probabilità di realizzazione. Ma questo fa parte di un problema piú generale non ancora completamente risolto, che sarà trattato nella prossima sezione.

7. *Caratteri essenziali della risoluzione della dicotomia reversibilità/irreversibilità su basi probabilistiche e loro connessioni con le attuali linee di ricerca.*

In questo paragrafo conclusivo si tenterà una ricapitolazione logica complessiva di tutti gli elementi che confluiscono nella linea di risoluzione della dicotomia reversibilità/irreversibilità secondo un'interpretazione probabilistica, della quale sono già stati dati i lineamenti essenziali di carattere fisico e storico nei paragrafi precedenti. Questo permetterà anche una breve illustrazione di alcune delle ricerche attualmente perseguite con metodi matematici rigorosi su problemi connessi con i fondamenti della meccanica statistica dell'equilibrio e del non-equilibrio.

Tali considerazioni saranno basate in maniera determinante su molti dei concetti introdotti da Lanford [1976] nella sua analisi, condotta con tecniche matematiche rigorose, della derivazione dell'equazione di Boltzmann dalle equazioni della meccanica classica per sistemi con un numero molto grande di gradi di libertà.

Si prenda quindi in considerazione un sistema fisico per il quale possano essere applicati i metodi di descrizione fenomenologica del suo comportamento termodinamico, secondo i concetti introdotti nel § 2. Per fissare le idee si può considerare un gas diluito contenuto in un recipiente di volume molto grande.

Nelle condizioni piú generali, e quindi anche fuori dall'equilibrio termodinamico, lo stato macroscopico del sistema considerato è descritto da un certo numero di grandezze fisiche, variabili in generale da punto a punto e nel tempo, come per esempio la densità di massa $\rho(x, t)$, la densità di corrente $\mathbf{j}(\mathbf{x}, t)$, la densità di energia interna $\varepsilon(\mathbf{x}, t)$, la temperatura $T(x, t)$, ecc.

È importante porre esplicitamente in rilievo che tali grandezze fisiche sono naturalmente suscettibili di misurazione, almeno in linea di principio, mediante opportuni dispositivi sperimentali. La loro evoluzione nel tempo si svolge secondo le equazioni cinetiche, descritte nel § 2, come è possibile verificare, nei limiti degli errori sperimentali, mediante una serie di esperimenti opportunamente condotti.

È questo quindi il livello di descrizione macroscopica del sistema, caratterizzato da grandezze fisiche effettivamente misurabili e da leggi fenomenologiche di evoluzione determinate e verificate sperimentalmente. S'indicherà in generale con $A(t)$ lo stato macroscopico del sistema all'istante t, e con $\mathfrak{A}$ l'insieme delle possibili configurazioni macroscopiche. Sulla base delle equazioni cinetiche fenomenologiche il sistema, partendo da una certa configurazione iniziale $A(0)$ all'istante $t=0$, percorrerà una determinata traiettoria $A(t)$ nello spazio $\mathfrak{A}$. È stato posto esplicitamente in rilievo nel § 2 che le equazioni di evoluzione fenomenologiche non sono in generale invarianti per inversione temporale, il livello macroscopico è quindi caratterizzato da una evoluzione temporale di natura intrinsecamente irreversibile.

Poiché la determinazione macroscopica dello stato coinvolge una serie concreta di misure, sarà certamente possibile introdurre sullo spazio degli stati $\mathfrak{A}$ un concetto di distanza, che precisi di quanto sono differenti in generale due stati A_1 e A_2, in maniera tale che stati vicini siano molto simili dal punto di vista fisico.

Occorre anche ricordare che l'inevitabilità degli errori sperimentali connessi con una qualsiasi operazione di misura induce sullo spazio degli stati macroscopici $\mathfrak{A}$ una sorta di struttura «sfumata», per cui ogni affermazione fatta sugli stati macroscopici dovrà essere intesa come valida nei limiti degli errori sperimentali.

Per molte delle considerazioni fisiche che possono essere fatte sulla natura e sul comportamento del sistema questa descrizione di tipo fenomenologico è completamente sufficiente.

D'altra parte, sulla base dell'ipotesi molecolare validamente confermata dall'esperienza, è noto che i sistemi termodinamici sono costituiti da aggregati di molecole che interagiscono tra di loro mediante lo scambio di forze la cui natura può essere anche sperimentalmente determinata. Per cui si pone in termini estremamente naturali il problema di spiegare il complesso comportamento termodinamico dei sistemi fisici mediante la natura e il comportamento dei suoi costituenti elementari a livello molecolare.

Viene cosí introdotto un secondo livello, piú approfondito e dettagliato, di descrizione del sistema fisico, per il quale ora le grandezze rilevanti sono costituite da posizioni e velocità (e altri parametri interni se necessari) delle molecole che compongono il sistema, mentre si assume che la loro evoluzione possa essere descritta dalla meccanica classica hamiltoniana, sulla base delle leggi d'interazione tra le molecole supposte note. Naturalmente questo livello microscopico di descrizione è di piú difficile accesso dal punto di vista della determinazione sperimentale dei valori delle grandezze fisiche coinvolte.

S'indicherà in generale con $P(t)$ lo stato microscopico del sistema all'istante t, come spiegato nel § 1, e con Γ l'insieme delle possibili configurazioni microscopiche. Secondo le equazioni di Hamilton il punto rappresentativo del sistema $P(t)$ percorre in modo perfettamente reversibile un'orbita nello spazio delle fasi Γ.

Una delle caratteristiche piú importanti dei sistemi meccanici, che vengono assunti come modelli a livello cinetico-molecolare di concreti sistemi termodinamici, è costituita, come si è osservato precedentemente, dalla loro enorme complessità. Questo comporta in generale un'estrema instabilità delle orbite, per cui punti inizialmente vicini nello spazio delle fasi divergono sempre piú con il trascorrere del tempo, mentre regioni nello spazio delle fasi inizialmente regolari evolvono in regioni, dello stesso volume per il teorema di Liouville, di forme estremamente complesse sviluppando protuberanze filamentose che si avvolgono su se stesse. Al contrario le leggi di evoluzione macroscopiche mostrano una stabilità molto maggiore, in accordo peraltro con la loro determinazione fenomenologica. La precisa caratterizzazione in termini matematici delle proprietà di estrema instabilità delle orbite dei sistemi meccanici complessi costituisce oggetto di investigazione nell'ambito della teoria generale dei sistemi dinamici astratti e della moderna teoria ergodica [cfr. Arnol'd e Avez 1967].

La connessione tra la descrizione data dal livello macroscopico, direttamente accessibile dal punto di vista termodinamico alle misure e alle verifiche sperimentali, e quella offerta dal livello microscopico, piú strettamente legata alle proprietà elementari del sistema fisico, può essere ottenuta mediante appropriati procedimenti di media spaziale e temporale (cfr. § 5). Quindi, assegnata una configurazione microscopica P, sarà possibile determinare univocamente, almeno in linea di principio, la corrispondente configurazione macroscopica A applicando ripetutamente questi opportuni procedimenti di media. Ovviamente, nel corso di questi procedimenti, molti dei dettagli della descrizione microscopica vengono persi, per cui si ottiene una sostanziale "contrazione" della descrizione nel passaggio dal livello microscopico elementare a quello macroscopico fenomenologico.

Questa contrazione implica in particolare che diversi stati microscopici, anche notevolmente diversi tra di loro, possono corrispondere allo stesso stato macroscopico. Se in piú si tiene anche conto dell'indeterminazione, dovuta agli errori sperimentali di misura, con cui può essere conosciuta una data configurazione macroscopica, si può arrivare a precisare il concetto molto importante di «stella» di stati microscopici $\Gamma(A)$ associata allo stato macroscopico A (alcuni dei concetti introdotti e la relativa terminologia sono in parte mutuati dall'approfondita trattazione critica dei fondamenti della meccanica statistica data dagli Ehrenfest [1912]). Nella stella $\Gamma(A)$, che costituisce un sottoinsieme dello spazio delle fasi Γ, si pongano per definizione tutte le situazioni dinamiche microscopiche che, tramite gli appropriati procedimenti di media spaziale e temporale, dànno luogo a stati macroscopici che differiscono dallo stato assegnato A nei limiti della tolleranza sperimentale permessa.

Per ottenere la completa risoluzione, su basi probabilistiche, della dicotomia tra il carattere reversibile dell'evoluzione a livello microscopico e quello intrinsecamente irreversibile del livello macroscopico, occorrerebbe poter dimostrare in termini rigorosi il seguente enunciato: «Si supponga di considerare uno stato macroscopico iniziale A e la corrispondente stella di stati microscopici $\Gamma(A)$. Sia $A(t)$ lo stato evoluto da A all'istante t secondo le equazioni fenomenologiche irreversibili. Sia $P(t)$ lo stato microscopico ottenuto per l'istante t tramite l'evoluzione hamiltoniana reversibile a partire da un qualsiasi stato P appartenente alla stella $\Gamma(A)$ nell'istante iniziale. Allora per la maggior parte degli stati iniziali P scelti in $\Gamma(A)$ si ha che $P(t)$ appartiene alla stella $\Gamma(A(t))$».

È chiaro che queste affermazioni traducono l'essenza della soluzione data da Boltzmann al problema dell'interpretazione su basi microscopiche delle leggi della termodinamica fenomenologica. Infatti, se l'enunciato fosse vero, nonostante l'evoluzione a livello microscopico avvenga in modo reversibile, essa apparirebbe svolgersi secondo le leggi cinetiche irreversibili se osservata a livello macroscopico tramite la misura delle variabili fenomenologiche locali, purché si escludano alcune delle possibili traiettorie microscopiche considerate scarsamente probabili.

Purtroppo lo stato attuale delle ricerche nel campo della dinamica qualitativa dei sistemi con un numero molto grande di gradi di libertà non permette di dare ancora né una dimostrazione rigorosa delle affermazioni fatte ma nemmeno una specificazione completa del loro significato, specialmente dove si viene a precisare il concetto secondo cui «con grande probabilità» uno stato iniziale P scelto in $\Gamma(A)$ si evolve in uno stato $P(t)$ che sia in $\Gamma(A(t))$.

D'altra parte i recenti risultati di Lanford [1976] costituiscono una prova soddisfacente della sostanziale correttezza della soluzione delineata, in quanto permettono di ricavare in termini rigorosi dalla meccanica reversibile dei sistemi con un numero molto grande di gradi di libertà la validità in termini probabilistici dell'equazione irreversibile di Boltzmann. I limiti della presente trattazione non consentono di dare un cenno anche solo schematico dei contenuti e dei metodi di questa importante analisi; ci si limita ad osservare che essa non costituisce una completa soluzione del problema solo perché, per limitazioni di tipo tecnico dovute ai metodi di dimostrazione utilizzati, la validità dell'equazione di Boltzmann è assicurata solo per un intervallo di tempo piccolissimo a partire dall'istante iniziale. Per poter quindi ottenere risultati piú soddisfacenti dal punto di vista fisico sono necessari metodi tecnicamente piú potenti.

Attualmente anche la scuola dei probabilisti sovietici (tra cui Dobrušin e Sinaj) è attivamente coinvolta nello studio rigoroso della dinamica dei sistemi con un numero molto grande di gradi di libertà (al limite infinito), nell'ambito del problema della connessione della meccanica con la termodinamica. Anche in Italia, specialmente ad opera del gruppo di Roma condotto da Giovanni Gallavotti, ricerche di questo tipo sono state perseguite con successo negli ultimi anni.

Uno dei problemi piú importanti che occorre affrontare, anche solo per dare un significato ben preciso all'enunciato introdotto precedentemente, è quello di specificare il senso con cui le asserzioni di tipo probabilistico debbono essere intese. In particolare che significato deve essere dato all'affermazione «Per la maggior parte degli stati iniziali P scelti in $\Gamma(A)$»? Forse uno dei contributi piú rilevanti offerti dall'analisi di Lanford è stato proprio quello di dare una specificazione precisa, valida nei limiti del modello da lui considerato, del significato delle affermazioni di tipo probabilistico. Per una soddisfacente soluzione generale di questo problema sarebbe necessaria l'introduzione di un'opportuna misura μ sullo spazio delle fasi Γ in maniera tale che ogni regione Γ' di Γ verrebbe ad acquistare un «peso», dato dalla sua misura $\mu(\Gamma')$, indicativo della probabilità di realizzazione delle situazioni microscopiche contenute in Γ'. Una volta in possesso di tale misura a priori la validità in senso probabilistico delle equazioni cinetiche di tipo fenomenologico assumerebbe il seguente significato. Si considerino uno stato iniziale macroscopico A e la corrispondente stella di punti $\Gamma(A)$, per un determinato t si consideri quella parte $\Gamma'(A)$ di $\Gamma(A)$ fatta dai punti P tali che i loro evoluti $P(t)$ siano in $\Gamma(A(t))$. Il rapporto $\mu(\Gamma'(A))/\mu(\Gamma(A))$ può essere assunto come valutazione quantitativa della frazione di punti aventi buon comportamento termodinamico. Se fosse possibile dimostrare che il valore di questo rapporto è molto vicino a 1, ne risulterebbe una precisa caratterizzazione della validità in senso probabilistico delle equa-

zioni cinetiche. I punti esclusi, dotati di cattivo comportamento termodinamico, avrebbero scarsa probabilità di verificarsi. È chiaro peraltro che al crescere di t la porzione di $\Gamma(A)$ fatta di punti con cattivo comportamento termodinamico è destinata a crescere, perché per un sistema finito comunque grande vale il teorema di ricorrenza di Poincaré che alla lunga deve invalidare le equazioni irreversibili. Quindi un buon comportamento termodinamico si può ottenere solo nel limite in cui il numero di gradi di libertà del sistema (e quindi il numero delle particelle che lo costituiscono) diviene infinito. Questo giustifica il grande interesse che ha recentemente sollevato lo studio delle proprietà meccaniche di sistemi con un numero infinito di gradi di libertà, per i quali sono stati dimostrati importanti teoremi di esistenza e di unicità ed è stato analizzato il comportamento qualitativo delle soluzioni, nel corso di ricerche svolte negli ultimi anni prevalentemente negli Stati Uniti e nell'Unione Sovietica, ma anche in altri paesi tra cui l'Italia.

È anche chiaro che questo problema della misura a priori sullo spazio delle fasi è strettamente connesso con il problema ergodico, di cui sono stati dati dei rapidi cenni nel § 4. Anzi, l'osservazione secondo cui un buon comportamento termodinamico può ottenersi solo nel limite di sistemi con un numero infinito di gradi di libertà ha enormemente stimolato un rinnovato interesse verso il problema ergodico esteso ora anche ai sistemi infiniti.

Occorre però anche osservare che il problema della misura a priori sullo spazio delle fasi, che dovrebbe permettere una precisazione quantitativa delle affermazioni di natura probabilistica, coinvolge anche delle considerazioni di natura strettamente fisica sul grado di realizzabilità e controllabilità di determinate situazioni microscopiche. A tale proposito possono essere utili le seguenti considerazioni. Si supponga di partire da una determinata configurazione macroscopica A e si scelga un qualsiasi stato microscopico P in $\Gamma(A)$ di buon comportamento termodinamico, tale quindi che il suo evoluto $P(t)$ sia in $\Gamma(A(t))$. All'istante t si effettua una inversione temporale su $P(t)$ e su $A(t)$, ottenendo rispettivamente $P'(t)$ e $A'(t)$ che verranno assunti come nuovi stati iniziali. È chiaro che P' è nella stella $\Gamma(A')$; però, dopo un intervallo di tempo t, lo stato P' sarà ritornato in P a causa dell'evoluzione reversibile, mentre A' *non* ritorna in A poiché le leggi di evoluzione macroscopica sono irreversibili. Si ottiene quindi che P', pur essendo in $\Gamma(A')$, *non* ha buon comportamento termodinamico. Tra le proprietà fisiche della configurazione microscopica P' che impediscono un buon comportamento termodinamico svolge probabilmente un ruolo rilevante l'esistenza di precise correlazioni a lunga distanza tra i gradi di libertà microscopici di P', ereditate dalla configurazione iniziale P tramite l'evoluzione reversibile. Queste correlazioni a lunga distanza sono di difficile controllo macroscopico, per cui ogni tentativo di riprodurre P' genererà degli stati P'' vicini a P' e ancora nella stella $\Gamma(A')$, però di buon comportamento termodinamico. Come conclusione prevale sempre la dinamica macroscopica irreversibile, mentre la natura fisica dell'irreversibilità si viene a precisare, secondo la lucida analisi data da Prigogine [1962], nella naturale tendenza delle eccitazioni localizzate e rilevabili macroscopicamente, che caratterizzano fisicamente le deviazioni dallo stato di equilibrio termodinamico, a disperdersi nei mezzi materiali di dimensioni infinite in eccitazioni sempre piú complesse e diffuse e quindi meno rilevabili macroscopicamente.

È in questo modo che una dinamica sostanzialmente reversibile, ma enormemente complessa, a livello microscopico appare acquistare caratteri di irreversibilità nel passaggio alla descrizione macroscopica.

Si può quindi concludere, in pieno accordo con i risultati di una linea ininterrotta di ricerche, che partono dall'impostazione originaria di Boltzmann e si sviluppano nelle loro implicazioni fisiche e matematiche fino ad includere le piú recenti acquisizioni della fisica teorica moderna, che la proposta di risoluzione della dicotomia reversibilità/irreversibilità su basi probabilistiche si è dimostrata estremamente ricca di continui sviluppi, che non solo hanno condotto ad una sostanziale progressiva chiarificazione dei punti cruciali della proposta stessa, ma sono stati anche di stimolo per ulteriori importanti investigazioni sui fondamenti della meccanica statistica, sulla dinamica qualitativa dei sistemi infiniti e in generale sulle relazioni tra meccanica e probabilità.

Il fatto che molti degli aspetti fisici e matematici del problema non abbiano ancora trovato una sistemazione definitiva, ma siano tuttora aperti, non fa che accrescerne l'interesse anche dal punto di vista della ricerca contemporanea. [F. G.].

Arnol'd, V. I., e Avez, A.
1967 *Problèmes ergodiques de la mécanique classique*, Gauthier-Villars, Paris.

Bogoljubov, N. N.
1946 *Problemy dinamičeskoj teorii v statističeskoj fisike*, Moskva-Leningrad (trad. ingl. North Holland, Amsterdam 1962).

Boltzmann, L.
1896-98 *Vorlesungen über Gastheorie*, Barth, Leipzig.

Born, M., e Green, H. S.
1949 *A General Kinetic Theory of Liquids*, Cambridge University Press, Cambridge.

Ehrenfest, P., e Ehrenfest, T.
1911 *Begriffliche Grundlagen der statistichen Auffassung in der Mechanik*, in *Encyklopädie der mathematischen Wissenschaften*, IV. *Mechanik*, sez. D, n. 32, Teubner, Leipzig, pp. 1-90.

Gibbs, J. W.
1902 *Elementary Principles in Statistical Mechanics*, Scribner, New York; ora in *Collected Works*, vol. II, parte I, Yale University Press, New Haven Conn. 1948.

Grad, H.
1958 *Principles of the Kinetic Theory of Gases*, in *Handbuch der Physik*, XII. *Thermodynamics of Gases*, Springer, Berlin, pp. 205-94.

Kirkwood, J. G.
1946 *The statistical mechanical theory of transport processes*, I. *General theory*, in «Journal of Chemical Physics», XIV, 3, pp. 180-201.
1947 *The statistical mechanical theory of transport processes*, II. *Transport in gases*, in «Journal of Chemical Physics», XV, 1, pp. 72-76.

Lanford III, O. E.
1976 *On a derivation of the Boltzmann equation*, in «Astérisque», n. 40, pp. 117-37.

Lebowitz, J. L., e Penrose, O.
1973 *Modern ergodic theory*, in «Physics Today», XXVI, 2, pp. 23-29.
Prigogine, I.
1962 *Non-Equilibrium Statistical Mechanics*, Interscience, New York.
Prigogine, I., e Nicolis, G.
1977 *Self-Organization in Non-Equilibrium Systems*, Wiley, New York.
Prigogine, I., e Stengers, I.
1978 «Energia», in *Enciclopedia*, vol. V, Einaudi, Torino, pp. 412-38.
Ruelle, D.
1969 *Statistical Mechanics: Rigorous Results*, Benjamin, New York.
Yvon, J.
1935 *La théorie statistique des fluides et de l'équation d'état*, Hermann, Paris.

La nostra stessa esperienza di **vita** (cfr. **empiria/esperienza**) mostra come non vi sia **simmetria** tra il passato (cfr. **passato/presente**) e il **futuro**. Ricordiamo il passato (cfr. **memoria**) ma non conosciamo il futuro, pur cercando di prevederlo (cfr. **anticipazione, previsione e possibilità**). Allo stesso modo, in **fisica**, il principio della crescita dell'**entropia** impone un'irreversibilità del tempo (cfr. **tempo/temporalità**); tuttavia, lo studio di **un corpo** in movimento (cfr. **moto**) mostra come in questo caso il tempo (cfr. **spazio-tempo**) sia reversibile. Questo disaccordo fra la meccanica e la termodinamica (cfr. **energia, equilibrio/squilibrio**) pone un problema che fisici e matematici (cfr. **matematiche**) si sforzano di risolvere, in particolare attraverso l'uso di metodi (cfr. **metodo**) statistici (cfr. **decisione, giochi, induzione statistica, probabilità**).

Stato fisico

1. *La nozione intuitiva.*

Da un punto di vista euristico per stato di un sistema fisico s'intende un particolare modo di essere del sistema (come, per esempio, quando si parla di stato solido, stato legato, stato di equilibrio) o, piú specificamente, un complesso di caratteristiche osservabili e mutuamente indipendenti del sistema che ne determinano completamente il comportamento. Poiché termini come 'sistema', 'comportamento' del sistema, 'descrizione' di tale comportamento, assumono un significato preciso soltanto nell'ambito di una specifica teoria, lo stesso accadrà per il concetto di stato; e ciò giustifica la pluralità degli usi, con significati diversi, di questo termine in contesti diversi (stato meccanico, stato termodinamico, stato quantistico). A questa pluralità di usi tecnici si contrappone tuttavia una ben radicata intuizione qualitativa secondo cui lo stato di un sistema, definendo un «modo di essere» del sistema stesso, è un suo attributo intrinseco e perciò non può dipendere dal rapporto soggetto conoscente / sistema. Secondo una tale concezione la molteplicità degli usi tecnici della parola 'stato' in differenti contesti è dovuta unicamente al fatto che alcune teorie corrispondono a una descrizione incompleta del sistema in questione, una descrizione completa essendo sempre, almeno in linea di principio, possibile. Esisterebbe quindi un unico concetto di stato, cioè quello relativo al livello piú fine della descrizione del sistema, tutti gli altri sarebbero versioni approssimate di questo o, in ogni caso, ad esso riducibili. Questa intuizione, associata al concetto di stato, è strettamente legata alla concezione meccanicistica della natura della quale costituisce uno dei presupposti impliciti. È interessante, a questo proposito, notare che Newton, nei *Principia*, non include il concetto di stato nella stringente analisi logica a cui sottopone la massima parte dei concetti che sono a fondamento della meccanica, e che lo stesso avverrà con Mach nella sua analisi storico-critica della meccanica newtoniana. Questa mancanza non è da attribuirsi a superficialità ma solo al fatto che la definizione che emerge dalla meccanica newtoniana del concetto di stato si presenta come un affinamento e una precisazione dell'idea, legata al senso comune, di stato come insieme completo di attributi di un sistema e perciò a questo livello non sembra meritare particolari approfondimenti.

La relatività e la meccanica quantistica mettono in crisi l'intuizione tradizionale associata a termini come 'sistema', 'descrizione completa', 'proprietà', ... e impongono, per le esigenze di sviluppo del pensiero scientifico, prima ancora che per esigenze epistemologiche e filosofiche, un esame molto piú approfondito non solo del concetto di stato di un sistema ma anche dello stesso concetto di sistema fisico.

Prima di discutere questi sviluppi sarà bene analizzare esplicitamente il concetto di stato fisico cosí come emerge dall'evoluzione di quella concezione della natura che è stata designata col termine, un po' ambiguo, di 'meccanicismo'.

L'intuizione associata a questo concetto è radicata nella nostra cultura da una tradizione piú che secolare e gli *idola fori* sviluppatisi intorno ad essa sono ancor oggi fonte di confusione nel dibattito sulla teoria quantistica (specialmente sulla teoria della misura quantistica).

La relazione tra il concetto «intuitivo», che un determinato individuo (o gruppo omogeneo di individui) può associare al termine 'stato', e la sua definizione generale è gravida di tutte le ambiguità relative al rapporto tra pensiero e linguaggio.

La definizione, intesa come momento in cui un concetto viene tradotto in termini linguistici, è piú astratta e quindi piú povera dell'intuizione e in questo senso allontana dalla realtà immediata delle percezioni e delle immagini. Ma talvolta questa maggiore povertà si tramuta in un'operazione di sfrondamento, di liberazione da pregiudizi e limitazioni che l'uso e le condizioni storiche avevano accumulato su un concetto, in una base di partenza per lo sviluppo di una nuova intuizione relativa a tale concetto; in questo senso l'astrazione avvicina alla realtà.

Nel caso delle teorie fisiche il momento della definizione è quello in cui viene costruito un modello matematico. Ma affinché un modello matematico sia significativo (non solo nel senso di contenere informazione, ma anche in quello di trasmetterla) occorre stabilire una corrispondenza tra entità matematiche ed entità riguardanti il mondo naturale, e il tramite di questa corrispondenza è il linguaggio comune. Attraverso questo può accadere che, nell'epoca di transizione da una teoria a un'altra, il vecchio intuito si intersechi col nuovo e si sovrapponga in parte al nuovo modello. Una situazione del genere si è verificata con il concetto di «stato» in meccanica quantistica, ed è stata sorgente di confusione nel dibattito sui fondamenti concettuali e sull'interpretazione della teoria quantistica.

Le ambiguità sul concetto di stato nascono dal fatto che esso è fondato su nozioni non precisamente definite come quelle di «sistema» o di «individualità»: solo con un atto teorico viene deciso che diversi complessi di percezioni si riferiscono a diversi modi di essere dello stesso sistema e non a diversi sistemi. Nel caso di oggetti macroscopici l'atto di isolare sistemi individuali dal complesso delle nostre percezioni è connaturato al nostro modo d'interagire col mondo esterno; la situazione cambia nel caso di oggetti microscopici: nella fisica delle particelle non è affatto banale definire quali complessi di percezioni si riferiscono a sistemi «diversi» e quali a diversi modi di essere (stati) dello «stesso» sistema. Per esempio, nella fisica delle interazioni forti, il protone e il neutrone vengono considerati come due «stati» di una stessa particella: il «nucleone»; mentre se s'introducono altri tipi d'interazione (debole, elettromagnetica, ecc.) essi vengono considerati particelle diverse.

Nella pluralità di accezioni del termine 'stato' è tuttavia possibile isolare alcune fondamentali caratteristiche comuni che permettono di unificare le varie definizioni (cfr. § 9). Ma l'unificazione a livello di definizione formale non riflette un'unità delle intuizioni associate a questo termine; queste infatti dipen-

dono dalla concezione che uno ha di che cosa significa descrivere un sistema nel senso di una scienza sperimentale. Da questo punto di vista i concetti di stato come emergono dalle teorie che sono state chiamate «meccanicistiche», e dalle teorie quantistiche, sono completamente diversi.

2. *Il concetto di stato in una teoria meccanicistica.*

In generale una teoria fisica è definita da un insieme $\mathcal{O}$ di grandezze osservabili e da un insieme $\mathcal{A}$ di affermazioni che riguardano i valori assunti da tali grandezze. Queste affermazioni possono essere di due tipi: 1) affermazioni che definiscono o limitano l'insieme dei valori a priori possibili di una singola grandezza (per esempio il fatto che lo spin può assumere solo valori interi o semiinteri, o che nessuna velocità può essere superiore a quella della luce nel vuoto); 2) affermazioni ipotetico-deduttive, cioè quelle che da alcune informazioni sui valori assunti da certe grandezze permettono di estrarre delle informazioni sui valori assunti da altre.

Tutte le affermazioni del primo tipo vanno considerate come particolari leggi fisiche; altri esempi di leggi fisiche sono forniti da schemi di affermazioni del secondo tipo (la legge di gravitazione di Newton $F = -G(mM/r^2)(\mathbf{r}/r)$, per esempio, è uno schema di affermazioni ipotetico-deduttive, nel senso che comunque si fissino i valori di m, M e $\mathbf{r}$ si ottiene un'affermazione di questo tipo).

Nel seguito si distinguerà fra teoria fisica e modello matematico di questa, e si riterranno equivalenti due modelli se, indipendentemente dalle loro strutture matematiche – che possono essere molto diverse –, essi conducono allo stesso insieme di affermazioni fisiche.

Il concetto di grandezza osservabile sarà considerato come concetto primitivo (euristicamente, esso include qualsiasi cosa possa essere misurata, e i risultati delle misure su un'osservabile – i suoi valori – sono numeri reali).

Una osservabile A sarà detta funzione di un'osservabile B se esiste una funzione f, che a un valore di B associa un valore di A, tale che ogni volta che il risultato della misura di B è b, il risultato della misura di A è $a = f(b)$; in questo caso si scriverà $A = f(B)$.

Le osservabili di un insieme saranno dette funzionalmente indipendenti se nessuna di esse è funzione delle altre; spesso, quando non c'è possibilità di confusione col concetto d'indipendenza statistica, si dirà semplicemente che queste osservabili sono indipendenti. Nel seguito i simboli A, B, C, ... denoteranno dei complessi di grandezze osservabili relativi a un dato istante di tempo (per esempio $A = (q_t, p_t)$ = posizione e impulso di una particella all'istante t), e i loro valori in quell'istante saranno rispettivamente indicati con i simboli a, b, c, ... Quando si vorrà sottolineare la dipendenza temporale si scriverà A_t invece di A, e si userà la notazione $B = b$ per indicare il fatto che la misura del complesso di grandezze $B = (B_1, B_2, ...)$ ha dato i risultati $b = (b_1, b_2, ...)$ Ciò premesso, si possono enunciare i postulati generali che caratterizzano la concezione meccanicistica della natura:

POSTULATO I). *Esiste un insieme B (finito o infinito) di osservabili indipendenti tali che per ogni istante t e per ogni osservabile A, i valori assunti da A all'istante t sono funzione dei valori assunti da B all'istante t.*

POSTULATO II). *Se B è un insieme di osservabili e A è un'osservabile tale che, a un certo istante t, $A_t = f(B_t)$, allora in ogni altro istante s si ha $A_s = f(B_s)$.*

POSTULATO III). *Se B è un insieme di osservabili che soddisfa la condizione* I), *allora per ogni coppia di istanti $s < t$ esiste una funzione $T^B_{t,s}$ tale che $B_t = T^B_{t,s}(B_s)$.*

POSTULATO IV). *È possibile, in linea di principio, misurare i valori esatti, a un dato istante, di un qualsiasi insieme di grandezze e in un dato istante una grandezza non può assumere due valori diversi.*

POSTULATO V). *Sia le dipendenze funzionali di cui al punto* II) *sia la legge di evoluzione di cui al punto* III) *sono accessibili alla conoscenza umana mediante la ragione e gli esperimenti.*

Nel seguito ogni insieme di grandezze osservabili, che soddisfi la condizione del postulato I), sarà detto un insieme completo di osservabili. Con questa notazione si può riformulare il postulato I) affermando che esistono insiemi completi di osservabili. Ogni teoria che soddisfa tutti e cinque questi postulati sarà detta una teoria meccanicistica. Un'analisi, anche sommaria, del significato e della portata di questi postulati trascende i limiti del presente articolo; ci si limiterà perciò ad alcuni brevi commenti.

Il postulato I) garantisce che ogni teoria meccanicistica è una teoria esatta; esso è fisicamente privo di contenuto se non integrato dal postulato IV) (postulato di compatibilità). Nel seguito si vedrà che la teoria quantistica rifiuterà il postulato IV) in due modi: da una parte, affermando l'esistenza di insiemi di osservabili i cui valori esatti non sono simultaneamente misurabili (forma debole del principio d'indeterminazione di Heisenberg); dall'altra, affermando l'esistenza di osservabili i cui valori esatti a un dato istante non sono affatto misurabili.

Il postulato II) garantisce l'omogeneità temporale dei vincoli cinematici; in particolare da esso e dal postulato III) segue che la conoscenza di B all'istante s e della relazione funzionale $A_s = f(B_s)$ determina la conoscenza dei valori di A_t per ogni $t > s$.

Il postulato III) esprime il carattere deterministico dell'evoluzione temporale e infine il postulato V) (o postulato dell'accessibilità) esprime una generica affermazione di fede che non dovrebbe trovar posto in una trattazione rigorosamente formale, ma che è di fondamentale importanza per capire lo spirito della visione meccanicistica. Per Einstein, che di questa visione è stato l'ultimo grande rappresentante, la tenace fede nell'accessibilità è stata uno dei principali motivi che lo hanno spinto a rifiutare fino alla fine la nuova visione del mondo prospettata dalla teoria quantistica.

Si osservi infine che, per chiarezza, i postulati del meccanicismo sono stati espressi in termini non relativistici, ma non è difficile riformularli in modo tale che essi siano soddisfatti anche dalla teoria della relatività (sia ristretta sia generale).

Come si vede, i postulati del meccanicismo si possono suddividere in due

gruppi: uno contenente i postulati I), IV), V), che riguarda la situazione del sistema a un istante fissato; un altro, contenente i postulati II), III), che riguarda il divenire del sistema.

Dal primo gruppo di assiomi emerge un concetto di stato, come fotografia istantanea presa sul divenire del reale, che può essere considerato come la precisazione dell'idea di stato proposta da Bergson nell'*Evolution créatrice* con la sua analisi del «meccanismo cinematografico del pensiero».

Questo concetto di stato è associato a una concezione della natura secondo la quale:

1) Ogni «proprietà» di un sistema è specificata dal fatto che certe grandezze osservabili associate al sistema assumono dei valori ben precisi.
2) In ogni istante, tutte le proprietà di un sistema sono simultaneamente specificate (cioè tutte le grandezze osservabili assumono un valore ben preciso).
3) La specificazione di una qualsiasi proprietà in un certo istante è indipendente dalla verifica di questa (cioè il valore assunto da una grandezza osservabile in un dato istante è indipendente dal fatto che tale valore venga sperimentalmente misurato).

Le condizioni 1), 2), 3) costituiscono l'essenza del cosiddetto «postulato del realismo», enunciato esplicitamente per la prima volta da Einstein, Podolsky e Rosen [1935]. Come si vedrà piú avanti, gli sviluppi della teoria quantistica indurranno alcuni autori a rifiutare questo postulato, indebolendo la condizione 1) e negando decisamente le altre due.

Secondo il postulato del realismo, lo stato di un sistema è quindi una sezione istantanea del suo divenire, cioè l'insieme delle specificazioni di *tutte* le sue proprietà a un dato istante. Nel seguito di questo paragrafo il termine 'stato di un sistema' sarà sempre usato nel senso appena specificato. Questa concezione dipende solo dal postulato del realismo e non dalla particolare teoria, purché naturalmente la teoria includa un numero sufficiente di osservabili e di affermazioni intorno ad esse (tante cioè da poter specificare *tutte* le proprietà a priori attribuibili al sistema). È questa la formulazione precisa dell'idea intuitiva di stato come insieme completo di attributi intrinseci del sistema alla quale si è fatto riferimento (§ 1).

In questa concezione occorre distinguere due aspetti: uno storico, in cui le «proprietà a priori attribuibili al sistema» sono fissate dall'assegnazione di un insieme di osservabili riferite a una ben precisa teoria. Uno metastorico, assoluto, secondo cui il postulato del realismo non riguarda una specifica teoria, ma la natura in sé, e perciò quando, in una data epoca, si parla delle «proprietà a priori attribuibili al sistema», ci si riferisce anche a quelle proprietà che, in quell'epoca, non sono ancora state individuate (ed è soltanto in questo secondo senso che il concetto di stato viene svincolato da ogni teoria e riflette solo la visione del mondo espressa dal postulato del realismo).

3. *Il modello matematico di stato in una teoria meccanicistica.*

Gli assiomi I), IV) e V) delle teorie meccanicistiche implicano, come si è visto, il postulato del realismo e lo storicizzano riferendolo a una particolare teoria. Essi però contengono un'ulteriore informazione che permette di ridurre drasticamente la quantità d'informazione necessaria per determinare lo stato di un sistema, e rende con ciò possibile la formulazione del modello matematico. Infatti da questi postulati segue che, per conoscere i valori assunti da tutte le grandezze osservabili di un sistema in un dato istante, non è necessario misurare tutte queste grandezze, ma è sufficiente misurare i valori di un qualsiasi insieme di osservabili che soddisfa la condizione del postulato I), cioè di un insieme completo di osservabili. Per comprendere a fondo la portata di questa riduzione e in che senso essa permetta il passaggio dal concetto di stato come nozione fisica al modello matematico, si comincerà a considerare una classe particolare di teorie meccanicistiche, che verranno chiamate newtoniane, caratterizzate dal fatto che in esse il postulato I) è rafforzato nel modo seguente:

POSTULATO I_N). *Esiste un insieme completo finito di osservabili.*

Gli attributi che definiscono il modo di essere, cioè lo stato di un sistema a un dato istante, possono essere, e sono in generale, infiniti. Il postulato I_N) implica però che questa infinità attuale possa essere ridotta a un insieme finito.

Sia $B=(B_1, ..., B_N)$ un tale insieme finito. L'assegnazione dei valori $b=(b_1, ..., b_N)$ assunti da B in un qualsiasi istante t determina completamente lo stato del sistema in quell'istante nel senso che, se A è una qualsiasi altra osservabile, esiste una funzione f tale che il valore di A all'istante t è dato da $f(b)$. Il rapporto tra lo stato del sistema all'istante t e la condizione $B_t=b$ è lo stesso che tra l'attuale e il potenziale. Se si pensa allo stato del sistema al tempo t come a una fotografia in bianco e nero ottenuta per accostamento di infiniti punti grigi con tonalità variabili, allora si può pensare alla condizione $B_t=b$ come a una macchina che permette, dato uno qualsiasi di questi punti, di riprodurlo, previa esecuzione di certe operazioni. Tutti i punti possono essere riprodotti, ma solo uno alla volta e, poiché in un tempo finito un uomo può eseguire solo un numero finito di operazioni, la fotografia non può mai essere ricomposta integralmente.

Ma se «lo stato in sé» di un sistema, inteso come sezione temporale del suo divenire, è inaccessibile nella sua attualità alla conoscenza umana, ben diversa è la situazione con quegli insiemi che contengono in sé, a livello virtuale, tutta l'informazione relativa a tale stato.

Se $B=(B_1, ..., B_N)$ è un insieme di osservabili che soddisfa la condizione del postulato I_N), per il postulato IV) è possibile una misura simultanea di B. Se $S(B_j)$ è l'insieme dei valori a priori possibili per la grandezza B_j ($j=1, ..., N$), i risultati a priori possibili delle misure simultanee su B saranno le N-uple $b=(b_1, ..., b_N)$, dove b_j appartiene a $S(B_j)$ ($j=1, ..., N$). Il fatto che le grandezze dell'insieme B siano funzionalmente indipendenti equivale a dire che

tutte le N-uple $b=(b_1, ..., b_N)$ con b_j in $S(B_j)$ $(j=1, ..., N)$ rappresentano risultati a priori possibili di misure simultanee di B. Si denoti con $S(B)$ l'insieme di tali N-uple. Poiché ogni osservabile assume valori reali, $S(B)$ è un sottoinsieme di uno spazio a N dimensioni (lo spazio euclideo N-dimensionale). Dalla discussione precedente risulta che ogni punto di $S(B)$, cioè ogni insieme di valori $b=(b_1, ..., b_N)$, determina univocamente uno stato del sistema e che, viceversa, ogni stato del sistema determina un unico punto di $S(B)$. Riferendosi a questa corrispondenza biunivoca, si dirà che $S(B)$ è lo spazio degli stati del sistema nella rappresentazione associata a B. Si vede quindi come il solo postulato I_N) permetta di costruire un modello matematico di spazio degli stati del sistema; il modello è semplice poiché in esso uno stato è rappresentato da una N-upla ordinata di numeri reali. Lo spazio $S(B)$ non dipende solo dal sistema che si sta descrivendo, ma anche da una nostra scelta arbitraria, cioè la scelta dell'insieme completo di osservabili B. Se C è un altro insieme completo di osservabili, per la completezza di B esiste una funzione $f: S(B) \to S(C)$ tale che per ogni c in $S(C)$, esiste un b in $S(B)$ tale che $f(b)=c$. Ciò significa che il risultato di una qualsiasi misura di C a un dato istante può essere ottenuto misurando B in quell'istante e applicando la funzione f al risultato di questa misura. Poiché il risultato di una misura di B determina completamente uno stato e lo stesso vale per un risultato di una misura su C, non è possibile avere $f(b_1)=f(b_2)=c$ se b_1 e b_2 sono punti diversi di $S(B)$, altrimenti allo stesso stato del sistema, cioè quello determinato da C, corrisponderebbero due valori diversi del complesso di grandezze B, il che è impossibile poiché per il postulato IV) in uno stesso istante una grandezza non può assumere due valori diversi. Ciò implica che la corrispondenza $f: S(B) \to S(C)$ è biunivoca e che la funzione che esprime B in funzione di C è la funzione inversa di f.

Ora, tra i numeri reali, c'è una nozione naturale di «vicinanza» (cioè esiste una topologia naturale sui numeri reali). Ciò permette di definire una nozione naturale di «vicinanza» tra stati di un sistema dicendo che due stati sono vicini, o differiscono di poco, nella rappresentazione associata a B, se i corrispondenti punti di $S(B)$ sono vicini nel senso della topologia naturale su $S(B)$. Il postulato della continuità fisica si esprime allora dicendo che il concetto di vicinanza tra stati è intrinseco, cioè non dipende dalla rappresentazione matematica. In altri termini, se C è un altro insieme di osservabili e b_1 e b_2 sono stati, nella rappresentazione associata a B, tali che $f(b_1)=c_1$ e $f(b_2)=c_2$, allora dal fatto che b_1 e b_2 differiscano di poco si deduce che anche c_1 e c_2 differiscono di poco. In termini matematici, ciò esprime la continuità della funzione $f: S(B) \to S(C)$ e poiché, scambiando i ruoli di B e C, la f viene cambiata nella sua inversa, anche la continuità della funzione $f^{-1}: S(C) \to S(B)$. A questo punto un profondo teorema matematico, intuito da Dedekind e dimostrato da Brouwer ed Enriques (l'invarianza della dimensione per omeomorfismi) permette in molti casi (cioè quando i valori assunti dalle osservabili coprono un intervallo aperto di numeri reali) di dedurre una conseguenza non banale dal postulato della continuità fisica, e cioè: se in una teoria meccanicistica di tipo newtoniano un insieme completo di osservabili è formato da N grandezze, allora *ogni altro* insieme completo sarà formato esattamente da N grandezze. Piú precisamente, la conclusione è che, se anche gli spazi degli stati $S(B)$ dipendono dalla scelta dell'insieme completo B, il loro tipo topologico è indipendente da questa scelta e perciò costituisce una caratteristica intrinseca del sistema. Si denoti con S questo tipo topologico. S è uno spazio astratto che può essere identificato (tecnicamente mediante omeomorfismi) a ciascuno degli spazi $S(B)$, dove B è un insieme completo di osservabili. I punti di questo spazio astratto rappresentano gli stati del sistema, nel senso che sono in corrispondenza biunivoca con essi; tali punti non dipendono ormai da alcuna scelta arbitraria, ma sono intrinsecamente associati ad esso. Si dirà che S è lo spazio degli stati del sistema. Il fatto che S sia uno spazio topologico significa che, sull'insieme amorfo dei possibili stati, rappresentati intuitivamente come sezioni virtuali sul divenire del sistema, è stata aggiunta una struttura teorica, cioè un concetto di vicinanza, di somiglianza tra stati, che già non riguarda il sistema in sé, ma il nostro rapporto con esso; questa ulteriore arbitrarietà si manifesterà piú chiaramente quando, per esprimere quantitativamente il concetto di «vicinanza» occorrerà introdurre una metrica e un'unità di misura. A queste arbitrarietà è legata la molteplicità dei modelli di spazi degli stati a priori compatibili con un dato sistema fisico. Il tentativo di eliminare queste arbitrarietà ha condotto, da una parte, allo studio di condizioni che fissino univocamente la struttura metrica sullo spazio degli stati in funzione di condizioni che, almeno in linea di principio, possono essere considerate sperimentalmente controllabili (come per esempio le equazioni di campo di Einstein in relatività generale); dall'altra, all'enunciazione di principî di invarianza che affermano l'equivalenza fisica di modelli che differiscono solo per diverse scelte (locali o globali) dell'unità di misura. Da quest'ultima problematica sono nate le «teorie di *gauge*» – creazione del matematico Weyl – che tanta importanza hanno assunto nella fisica contemporanea. L'identificazione di S con uno specifico $S(B)$, che è un sottoinsieme dello spazio euclideo N-dimensionale $\mathbf{R}^N$, si effettua mediante una particolare funzione (omeomorfismo) $\varphi_B: S \to \mathbf{R}^N$. In matematica, assegnare una tale funzione significa assegnare una famiglia di coordinate (globali) sullo spazio S. Dal punto di vista matematico, quindi, la scelta di un particolare insieme completo di osservabili per descrivere gli stati di un sistema, cioè di una rappresentazione, equivale alla scelta di una famiglia di coordinate (globali) sullo spazio degli stati di questo sistema. Il passaggio da una rappresentazione associata all'insieme completo B ad una associata all'insieme completo C corrisponde al cambiamento di coordinate definito dalla funzione $f_{CB} = \varphi_C \circ \varphi_B^{-1}: S(B) \to S(C)$.

Nella pratica spesso si fissa un sistema di coordinate, cioè una particolare rappresentazione in qualche modo privilegiata rispetto al problema che si sta trattando, e s'identifica lo spazio degli stati con lo spazio $S(B)$ associato a quella particolare rappresentazione. Cosí, per esempio, nella meccanica hamiltoniana dei sistemi con un numero finito di gradi di libertà, spesso lo spazio degli stati è identificato all'insieme delle coppie (posizioni-impulsi) mentre già nella termodinamica classica occorre frequentemente passare da una rappresentazione a un'altra e perciò si distingue tra il termine generico 'funzione di stato' e il ter-

mine specifico 'funzione di P, V, ...', designando il primo termine una generica grandezza termodinamica, e il secondo la forma esplicita della dipendenza funzionale che permette di esprimere i valori di questa grandezza in funzione dei valori assunti dall'insieme completo di osservabili (P, V, ...).

Un'altra astrazione da cui dipende la costruzione dello spazio degli stati come struttura matematica è quella di «insieme di tutti i valori a priori possibili» di una grandezza osservabile. In essa ci sono due tipi di estrapolazione rispetto ai risultati dell'esperienza: una prima consiste nel fatto che le misure disponibili dei valori di una grandezza sono sempre in numero finito, e perciò quando si parla di «tutti i valori possibili» di una grandezza si sta in pratica interpolando o estrapolando, sulla base di una intuizione storicamente datata, l'insieme dei dati sperimentali. Si è già dato il caso in cui alcune di queste estrapolazioni, apparentemente «ovvie», sono poi state smentite dagli esperimenti: ciò è avvenuto, per esempio, con la scoperta di un limite superiore per le velocità raggiungibili in natura (la velocità della luce nel vuoto), o con la scoperta del carattere discreto di alcune quantità tradizionalmente ritenute continue, come l'energia, il momento angolare in una direzione, ecc. Un'altra estrapolazione è legata al fatto che tutte le nostre percezioni hanno un carattere *locale*: ogni strumento fisico ha un raggio d'azione limitato, nel senso che esso permette di misurare solo i valori di una data osservabile che sono compresi entro un intervallo limitato. È possibile ampliare il raggio d'azione di uno strumento o costruirne un altro piú potente e, almeno teoricamente, questo processo può non avere fine. Ma si tratta di una potenzialità, non di un fatto attuale: è un fatto che esistono dei limiti precisi, legati a una data epoca, sulle misure effettuabili sui valori di qualsiasi grandezza osservabile.

Per tener conto di questi limiti nella costruzione del modello matematico è necessario considerare, invece dell'insieme $S(B)$ di tutti i valori di una grandezza B, gli insiemi di tutti questi valori compresi in un dato intervallo limitato. L'intervallo può variare e può essere arbitrariamente grande, ma sempre mantenendosi limitato.

A un insieme completo $B=(B_1, ..., B_N)$ di osservabili sarà quindi associato non piú un unico spazio degli stati $S(B)$, ma tutta una famiglia $U(B)$ di «spazi locali» definiti in modo analogo ad $S(B)$ ma con la differenza che, nelle N-uple $b=(b_1, ..., b_N)$, i valori di b_j possono variare soltanto all'interno di un intervallo limitato U_j contenuto in $S(B_j)$. Ciascun punto di $U(B)$ definisce univocamente uno stato del sistema, ma questa volta non è piú vero che *tutti* gli stati del sistema sono definiti da un punto di $U(B)$. Per il sottoinsieme degli stati del sistema individuato da $U(B)$ è possibile ripetere l'analisi, svolta precedentemente per tutti gli stati, e definire uno spazio topologico astratto U, i cui punti sono intrinsecamente associati a un certo sottoinsieme di stati del sistema e da cui la rappresentazione $U(B)$ – associata a un generico insieme completo di osservabili B – si deduce mediante introduzione di un sistema di coordinate, questa volta *locali*, cioè di una funzione (omeomorfismo) $\varphi_B : U \to \mathbf{R}^N$. Identificando i punti di ciascuno spazio astratto U con stati del sistema, e denotando con S l'insieme di tutti gli stati del sistema a priori possibili, le rappresentazioni matematiche di S saranno date non piú da identificazioni globali $\varphi : S \to \mathbf{R}^N$, ma da una famiglia $\{(U, \varphi)\}$ di coppie (U, φ), dove U è un sottoinsieme di S e φ è una identificazione di U con un sottoinsieme di $\mathbf{R}^N$. Si tratterà cioè di rappresentazioni locali, nel senso specificato sopra.

Nella terminologia matematica, una coppia (U, φ) con le proprietà elencate sopra è detta una carta locale. Se (U, φ) e (V, ψ) sono due carte locali, sull'insieme $U \cap V$ degli stati comuni ad U e V sono definite sia la rappresentazione associata a φ che quella associata a ψ. Il problema della loro connessione si discute esattamente allo stesso modo del problema, già analizzato, della connessione tra le rappresentazioni globali associate a due insiemi completi B e C, con l'unica differenza che S è sostituito da $U \cap V$.

Il risultato è analogo, e la funzione che definisce il passaggio dalle coordinate φ a quelle ψ è $\psi \circ \varphi^{-1}$. Se si accetta il postulato della continuità, questa funzione è un omeomorfismo. Poiché ogni stato è definito da almeno un insieme completo di osservabili, si ha che ogni punto di S appartiene ad almeno un U; in altri termini, l'insieme di tutti gli U è un ricoprimento di S. In matematica una struttura determinata da uno spazio topologico S, una famiglia $\{(U, \varphi)\}$ di carte locali su S tale che l'insieme di tutti gli U sia un ricoprimento di S e la funzione $\psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V)$ sia continua per ogni coppia di carte locali (U, φ), (V, ψ), è detta una varietà topologica.

Riassumendo quanto detto finora si ha che le condizioni 1) esistenza di un insieme finito completo di osservabili, cioè il postulato 1_N), 2) postulato di continuità fisica, 3) località delle nostre percezioni, 4) ipotesi che i valori assunti da una qualsiasi osservabile di un insieme completo costituiscono un intervallo aperto di numeri reali, permettono di determinare univocamente il modello matematico dello spazio degli stati. Tale modello risulta essere una varietà topologica. Se si rafforza il postulato di continuità fisica richiedendo non solo che a stati «vicini» in una rappresentazione corrispondano stati vicini in una qualsiasi altra, ma anche che rispetto a qualsiasi unità di misura la vicinanza tra due rappresentazioni dello stesso stato sia «dello stesso ordine di grandezza», si perviene al concetto di varietà differenziabile, che costituisce il modello abituale per lo spazio degli stati di una qualsiasi teoria meccanicistica di tipo newtoniano. Esempi tipici di queste teorie sono: la meccanica hamiltoniana dei sistemi con un numero finito di gradi di libertà; la termodinamica classica; la teoria generale dei sistemi dinamici a dimensione finita. Le caratteristiche specifiche della varietà (per esempio la dimensione) dipenderanno dal particolare sistema, ma la struttura generale del modello dipende unicamente dalle quattro condizioni elencate sopra.

Per le teorie meccanicistiche che non soddisfano il postulato 1_N) (per esempio la meccanica dei sistemi continui o, piú in generale, le teorie di campo), la linea concettuale del ragionamento precedente è ancora valida e ancora una volta suggerisce una varietà topologica (o differenziabile) come modello matematico dello spazio degli stati. Questa volta però la varietà non è modellata sullo spazio euclideo N-dimensionale, ma su uno spazio funzionale a infinite dimensioni. In questo caso anche se le difficoltà tecniche sono molto grandi, poiché la

teoria matematica degli spazi funzionali è molto meno sviluppata che non nel caso di spazi a dimensione finita, lo schema concettuale rimane essenzialmente lo stesso per quanto riguarda il modello matematico di spazio degli stati. Per quanto concerne le implicazioni sulla teoria della conoscenza invece il passaggio, nel modello matematico, da una varietà a dimensione finita a una a dimensione infinita comporta una notevole differenza.

Come si è visto, un punto nello spazio degli stati rappresenta lo stato del sistema e il numero delle sue coordinate rappresenta il numero minimo di misure simultanee indipendenti che bisogna eseguire per determinare completamente, nel senso virtuale spiegato prima, tutta l'informazione contenuta in quello stato.

In una varietà a dimensione infinita, tale numero è infinito e, poiché è impossibile eseguire simultaneamente una infinità attuale di misure, ne segue che in generale, se il sistema non è di tipo newtoniano, i suoi stati non sono accessibili all'esperimento nemmeno in quel senso virtuale che è stato descritto precedentemente. Per esempio, nella teoria dell'elettromagnetismo, i valori di tutte le osservabili rilevanti a un dato istante sono esprimibili come funzioni del campo elettromagnetico in quell'istante. Ma il campo elettromagnetico è definito dall'assegnazione di una coppia di vettori per ogni punto dello spazio. Anche volendo prescindere dai problemi coinvolti con la definizione del campo in un punto, anche volendo limitarsi a una zona limitata dello spazio, ci si trova davanti al problema di dover effettuare infinite misure simultanee per determinare il campo a un istante dato. La questione è molto meno pratica che teorica. Dal punto di vista dell'*uso* della teoria elettromagnetica spesso astute considerazioni di simmetria, ragionevoli approssimazioni, potenti mezzi di calcolo, o semplicemente l'uso accorto della teoria, suggeriscono precise indicazioni su come colmare le lacune dell'esperienza e conducono a successi che la ragione non riesce interamente a spiegare (Wigner parla di «irragionevole efficacia della matematica»). Inoltre, per una macchina finita, come il cervello umano, la differenza «in linea di principio» tra finito ma molto grande e infinito è irrilevante: per una tale macchina tutti i numeri da un certo punto in poi sono considerati, di fatto, infiniti. Tuttavia, da un punto di vista teorico, è indubbio che il concetto di stato in una teoria di campo classica, o piú in generale in una teoria meccanicistica non di tipo newtoniano, contiene a livello fondamentale piú elementi speculativi rispetto a una teoria di tipo newtoniano. Ancora non si è arrivati, come farà la teoria quantistica, a dichiarare l'inconoscibilità dello stato in senso meccanicistico, ma la conoscibilità non dipende piú solo dalla ragione e dall'esperimento, ma anche da intuizioni speculative che guidano la mente a colmare quelle lacune dell'esperienza che, per la prima volta, non hanno un carattere storico, ma di principio. Per conoscere lo stato dell'universo laplaciano basta partecipare a un barlume della conoscenza divina; per conoscere lo stato dell'universo einsteiniano occorre anche una scintilla di divina intuizione.

4. *Stato «meccanicistico» e stato «reale».*

L'analisi precedente mostra come dai postulati I), IV) e V) delle teorie meccanicistiche emerga un concetto intuitivo di stato di un sistema, come complesso di tutte le proprietà del sistema in un dato istante, e un preciso suggerimento sulla struttura atta a rappresentare, nel modello matematico, l'insieme astratto di tutti i possibili stati del sistema. D'altra parte, nell'intuizione associata a tutte le teorie meccanicistiche (tranne quelle che si limitano a priori alla descrizione di equilibrio, in cui l'evoluzione temporale è banale) l'idea di stato non esprime soltanto la bergsoniana fotografia istantanea dell'evoluzione temporale di un sistema, ma anche l'embrione da cui tale evoluzione si sviluppa. Questo aspetto è contenuto nei postulati II) e III), le cui conseguenze, ai fini della rappresentazione quantitativa del concetto di stato, saranno ora analizzate.

Dal postulato II) segue che le funzioni esprimenti i valori di un insieme completo di osservabili a un dato istante in funzione dei valori, nello stesso istante, di un altro insieme completo di osservabili, non dipendono dall'istante scelto. Poiché, secondo l'analisi del § 3, tali funzioni, con gli insiemi dei valori a priori possibili delle osservabili (che, per implicito postulato, neanche dipendono dal tempo), determinano univocamente il modello matematico, segue che non è necessario un modello di spazio degli stati per ogni istante di tempo, ed è possibile rappresentare l'evoluzione temporale del sistema come un punto, variabile nel tempo, in un unico spazio degli stati.

Nella rappresentazione (locale o globale) associata a un insieme completo di osservabili B, tale evoluzione sarà descritta dalla funzione $T^B_{t,s}$ definita dal postulato III) e, al variare di B tra tutti i possibili insiemi completi di osservabili, le funzioni $T^B_{t,s}$ determinano un'unica funzione $T_{t,s}$ globalmente definita sullo spazio degli stati S e che trasforma S in sé: infatti, se $(U(B), \varphi_B)$ e $(U(C), \varphi_C)$ sono carte locali su S associate rispettivamente agli insiemi completi B e C, i postulati I), II), III) implicano che su $U(B) \cap U(C)$ si deve avere

$$\varphi_B^{-1} \circ T^B_{t,s} \circ \varphi_B = \varphi_C^{-1} \circ T^C_{t,s} \circ \varphi_C.$$

Poiché $T_{t,s}$ è globalmente definita su S, qualunque sia lo stato σ_s del sistema al tempo s, il suo stato al tempo $t>s$ sarà dato da $T_{t,s}\sigma_s = \sigma_t$. In questo senso si dice che in una teoria deterministica il futuro (cioè lo stato del sistema a ogni istante $t>s$) è determinato dallo stato presente (cioè lo stato al tempo s). Si è già insistito sul fatto che lo stato σ nella sua globalità è un'astrazione teorica, non un concetto sperimentale. Analogamente la funzione $T_{t,s}$ è un'astrazione: per ogni previsione quantitativa, confrontabile con gli esperimenti, è necessario considerare la $T_{t,s}$ in qualche particolare rappresentazione, cioè in un particolare sistema di coordinate, che definiscono le grandezze sulle quali concretamente si faranno, e verificheranno sperimentalmente, delle previsioni. La funzione astratta $T_{t,s}$ invece si rivela spesso utile per considerazioni di tipo qualitativo sul sistema che coinvolgono proprietà indipendenti dalla rappresentazione, in particolare quelle che coinvolgono il comportamento del sistema in intervalli

temporali molto lunghi (o infiniti). Cosí per esempio, si dice che σ è uno stato stazionario (o di equilibrio) se, per ogni $s<t$, $T_{t,s}\sigma=\sigma$; che è uno stato di equilibrio stabile, se $\lim_{t\to+\infty} T_{t,s}\sigma'=\sigma$, per ogni stato σ' abbastanza vicino a σ, ...

Dai postulati II) e III) segue inoltre che, se $r<s<t$ sono arbitrari istanti di tempo, allora deve valere la relazione

$$(1) \qquad T_{t,r}=T_{t,s}\cdot T_{s,r}.$$

Questa relazione con l'aggiunta di ipotesi tecniche di regolarità (che non verranno enunciate), analoghe al postulato di continuità introdotto nel § 3, equivale ad affermare che l'evoluzione temporale degli stati di un sistema deterministico è retta da una equazione differenziale del prim'ordine. Nella rappresentazione associata a un insieme completo di osservabili B, questa equazione ha la forma

$$(2) \qquad \frac{d}{dt}B_t=X(B_t)$$

dove B_t è un vettore di uno spazio vettoriale reale (di dimensione finita se il sistema soddisfa il postulato I_N)).

Ogni equazione del tipo (2) descrive un'evoluzione deterministica del complesso di grandezze B, ma soltanto se B è un insieme completo di grandezze si può affermare che l'equazione (2) descrive l'evoluzione dello stato del sistema nel senso delle teorie meccanicistiche. Su questo punto si tornerà nel § 9.

Allo scopo di analizzare il rapporto tra lo «stato» definito da una concreta teoria meccanicistica e lo «stato reale», si cominci a osservare che ogni teoria meccanicistica è imperniata sui seguenti tre tipi di leggi:

1) leggi che riguardano i valori a priori possibili delle grandezze fisiche;
2) relazioni esatte tra i valori di differenti grandezze fisiche in uno stesso istante di tempo (quelle che si sono chiamate «vincoli cinematici»);
3) la legge di evoluzione di un insieme completo B di osservabili, cioè la forma esplicita dell'operatore X nell'equazione (2).

È importante osservare che, tra queste leggi, non ce n'è alcuna che permetta d'individuare un insieme completo di osservabili, o di stabilire se un dato insieme è completo oppure no.

In effetti, per ovvi motivi, un tale criterio di decisione, inteso come algoritmo sperimentale, non può esistere in linea di principio non appena la teoria ammetta un insieme infinito di osservabili che non siano banalmente date a priori come funzioni di un numero finito di esse. Come si vedrà piú in dettaglio nel paragrafo seguente, tutte le teorie di tipo meccanicistico, nel loro concreto sviluppo storico, risolvono questo problema semplicemente *decidendo* che un certo insieme di grandezze è un insieme completo.

Le motivazioni, che giustificano storicamente questa decisione, saranno analizzate nel paragrafo seguente. Ciò che per il momento interessa sottolineare è che, se la teoria vuole essere una teoria sperimentale, una tale arbitraria decisione è in linea di principio ineliminabile (infatti: se ci sono infinite osservabili non date a priori come funzioni di un insieme finito di esse, si è già visto che non si potrà mai decidere con un numero finito di esperimenti se un dato insieme di osservabili indipendenti è completo o no; d'altra parte, se si fissa a priori un numero finito di osservabili, non si potrà mai dimostrare sperimentalmente che queste definiscano *tutte* le proprietà fisiche del sistema). Come in ogni teoria matematica sufficientemente ricca esistono affermazioni che non possono essere né dimostrate né confutate all'interno della teoria, cosí, ora come in una qualsiasi epoca futura, l'affermazione che una data teoria meccanicistica fornisce una descrizione *completa* di un sistema, non può essere dimostrata all'interno di una qualsiasi teoria sperimentale.

La famosa affermazione di Lenin sulla «inesauribilità dell'elettrone» si può tradurre, nel nostro linguaggio, proprio come l'affermazione dell'impossibilità di dimostrare che un dato insieme di osservabili indipendenti relative al sistema-elettrone è, per questo, un insieme completo. Ma c'è di piú: non solo non si può sapere se una data descrizione è completa o no, ma non si può neppure sapere *quanto* si discosta dalla completezza. Ciò implica che nel programma del meccanicismo non solo non può essere inclusa la speranza di arrivare a una descrizione «fedele» (cioè completa) della natura, ma neppure quella di un'«approssimazione asintotica» (si veda l'articolo «Fisica» in questa stessa *Enciclopedia*) ad essa.

Se s'immagina lo «stato meccanicistico» di un sistema, inteso in senso assoluto come l'insieme di *tutte* le proprietà del sistema – esistenti obiettivamente e indipendentemente dal fatto che l'uomo le osservi o le misuri –, come un punto σ in uno spazio immaginario, si può allora pensare ai vari concetti «stato», relativi ai successivi sviluppi storici di una teoria meccanicistica, come a una successione di punti in questo spazio immaginario che «si avvicina sempre di piú» al punto σ, che rappresenta il *vero* stato. Ma anche se si accetta questa concezione per cosí dire «lineare» dello sviluppo della scienza, la cui inadeguatezza è stata messa in luce dalle analisi di Kuhn e di molti altri dopo di lui, non va confuso l'«avvicinarsi sempre di piú» (che può avvenire rimanendo sempre ben lontani dalla meta), con l'«approssimarsi asintotico» in cui è implicita l'idea che dopo un numero finito di passi la differenza tra approssimante e approssimato diventi trascurabile. Questo secondo tipo di approssimazione non può mai essere garantito nell'ambito di una teoria meccanicistica a causa della mancanza di un criterio obiettivo che permetta di valutare *quanto* lo «stato del sistema secondo una data teoria» si discosti dallo «stato reale» (s'incontra una situazione analoga quando si tenta di risolvere numericamente un'equazione: anche in questo caso, assegnare un metodo di approssimazione senza dare la stima dell'errore non fornisce alcuna informazione sulla soluzione).

In una qualsiasi teoria meccanicistica lo stato di un sistema è determinato da un insieme di caratteristiche obiettive, cioè i valori di un insieme completo di grandezze a un dato istante; ma cosa si debba intendere quando si parla di un «insieme completo» di osservabili lo *decide* la teoria. La stessa situazione si riproporrà nella teoria quantistica: la differenza sta solo nel *tipo* di decisione.

5. *Stato ed evoluzione.*

Il concetto di stato in una teoria meccanicistica è fondato, come si è visto, su quello di insieme completo di osservabili, e l'esistenza di tali insiemi è un postulato di queste teorie. Il passaggio dallo schema di teorie, definito dai postulati I)-V), a una specifica teoria e al corrispondente modello matematico, richiede l'assegnazione a priori, mediante un ulteriore postulato, di tutti i possibili insiemi completi di osservabili o, equivalentemente, di un insieme completo di osservabili e di tutti i «cambiamenti di coordinate» che consentono di passare dalla rappresentazione associata a quell'insieme alla rappresentazione associata a un qualunque altro insieme completo (cfr. § 3).

Ai nostri giorni, quando si parla di «postulati di una teoria fisica», s'intende semplicemente separare la parte puramente deduttiva della teoria da quella induttiva. I postulati sono cioè un complesso di affermazioni della teoria dal quale tutte le altre possono essere dedotte per via puramente logica. In questo senso essi rappresentano la sintesi concettuale delle conoscenze empiriche di una data epoca. Occorre però non dimenticare che la fisica non è una scienza deduttiva, non procede cioè per postulati e deduzioni, ma si serve di questi per coordinare e sviluppare i risultati ottenuti mediante esperimenti e induzioni. In particolare, una teoria fisica rappresenta la conclusione di un processo storico durante il quale i suoi contorni non sono nettamente delineati; non è ben chiaro quale sia «l'insieme di tutte le osservabili della teoria» né «l'insieme di tutte le affermazioni della teoria), ma si sa solo che alcune grandezze osservabili e alcune affermazioni su tali grandezze vengono ritenute «rilevanti» ai fini della teoria che si sta costruendo.

È in questa fase di sviluppo che emerge un legame, tra concetto di stato ed evoluzione, piú profondo di quello descritto nello sviluppo deduttivo della teoria. Infatti la possibilità di definire, in modo logicamente coerente, l'evoluzione di un sistema come l'evoluzione di un suo stato è un'astrazione a posteriori, possibile solo quando la teoria è già completa. Nel processo storico di formazione della teoria accade esattamente l'inverso, cioè è la scoperta di una legge di evoluzione ossia una equazione del tipo (2) per un certo complesso di grandezze B, che permette di definire cos'è un insieme completo di osservabili, quindi uno stato del sistema. Nella realtà storica, al contrario che nello sviluppo deduttivo, non si *cerca* l'evoluzione di un sistema completo di osservabili, ma si *definisce* completo un insieme di osservabili di cui si è scoperta la legge di evoluzione. L'insieme $\mathcal{O}$ di tutte le osservabili della teoria sarà allora per definizione l'insieme di tutte le funzioni dell'insieme B anzidetto (o meglio, tutte le funzioni che soddisfano certi requisiti di regolarità come funzioni reali definite su $S(B)$ = insieme dei valori delle grandezze B); e l'insieme di tutte le affermazioni della teoria sarà allora, sempre per definizione, l'insieme delle affermazioni del tipo: il valore dell'osservabile A – in $\mathcal{O}$ – all'istante t è $f(b_t)$.

Soltanto con Eulero e Lagrange il modello matematico dello stato meccanico, cosí come lo s'intende oggi, comincia ad emergere dalla constatazione che

l'equazione di Newton può essere scritta nella forma (2) se con B_t viene denotata la coppia posizione-velocità della particella all'istante t. Lo sviluppo di questa osservazione condurrà all'introduzione, da parte di Hamilton dello «spazio di fasi» associato a un sistema meccanico, che è il modello matematico per lo spazio degli stati di quel sistema. Su tale spazio è definita una particolare struttura (struttura simplettica) intimamente connessa alle evoluzioni di tipo newtoniano, nel senso che tutte le evoluzioni che «conservano questa struttura» sono descritte dalle equazioni di Hamilton, che sono una generalizzazione della legge di Newton.

In generale, il fatto che in una teoria meccanicistica lo spazio degli stati sia strettamente determinato dal tipo di legge di evoluzione (salvo, come si è già detto, nei casi in cui questa evoluzione è banale), si manifesta in addizionali strutture su questo spazio, oltre a quella generale di varietà differenziabile (cosí, per esempio, esistono dei campi la cui equazione di evoluzione è l'analogo, a infinite dimensioni, dell'equazione di Hamilton. Anche in questo caso sullo spazio degli stati – che è a dimensione infinita – è definita una struttura simplettica generalizzata).

Si può, in conclusione, affermare che ad ogni legge di evoluzione è possibile associare un concetto di «insieme completo di osservabili» (e quindi un concetto di stato) in modo tale che la risultante teoria soddisfi tutti i postulati del meccanicismo. Come è stato dimostrato, il meccanicismo come visione del mondo non è sottoponibile a verifica sperimentale; sottoponibili invece a tali verifiche sono i risultati dello studio sistematico delle equazioni che si presentano come utili modelli esatti di fenomeni naturali.

L'enorme successo storico di questo programma ha fatto sí che i postulati *impliciti* in tutti i modelli di tipo meccanicistico venissero *implicitamente* assunti al ruolo di verità universali della natura, e che da ciò che era un insieme di modelli si astraesse una ideologia scientifica.

Finora è stato analizzato il concetto di stato che di tale ideologia è l'espressione naturale. In quanto segue si vedrà come la crisi dell'ideologia meccanicista condurrà a un abbandono di questo concetto di stato. Il concetto che lo sostituirà è l'espressione non di una diversa teoria, ma di un nuovo modo di concepire la conoscenza della natura, perciò i contrasti che sorgeranno su tale concetto sono solo il riflesso dei piú profondi contrasti che hanno accompagnato l'abbandono della speranza di poter tradurre, in una teoria scientifica fondata sull'esperimento, la concezione meccanicistica della Natura.

6. *Stato nei sistemi aperti.*

Esiste tutto un complesso di teorie che, pur non situandosi, rispetto allo schema meccanicistico in una posizione di radicale rottura (come avverrà per la teoria quantistica) non rientrano in questo schema.

La principale differenza fra tali teorie e quelle meccanicistiche, nel senso del § 3, sta nel modo di porsi di fronte ai concetti di «sistema» e «stato di un siste-

ma». I sistemi che possono essere descritti dal meccanicismo sono quelli che contengono in sé l'embrione di tutto il loro sviluppo futuro, o quelli la cui evoluzione temporale può essere completamente prevista misurando i valori a un certo istante di un insieme di osservabili che riguardano *unicamente* il sistema stesso. Questi sono i cosiddetti «sistemi isolati» (o chiusi), quelli cioè sul cui comportamento l'ambiente esterno (rispetto a loro) non influisce in modo significativo (per esempio nel modello newtoniano il sistema solare viene considerato «isolato» dall'universo circostante). Esistono però in natura molti sistemi a cui è naturale attribuire una individualità, cioè che è naturale considerare come «sistemi», ma nel cui comportamento il ruolo dell'ambiente esterno non può essere considerato trascurabile; esempi tipici sono forniti dai sistemi biologici, dalle strutture ecologiche, dalle organizzazioni sociali o economiche. Tali sistemi sono chiamati «sistemi aperti». L'enorme varietà di sistemi che rientra in questa classe fa sí che siano stati elaborati per la loro descrizione vari tipi di modelli matematici. Una importante linea di pensiero descrive tali sistemi ispirandosi alle tecniche della termodinamica del non-equilibrio, e in questa direzione vanno ricordati gli importanti contributi di Prigogine e della sua scuola.

Un'altra linea di pensiero è piú direttamente ispirata a problemi provenienti dalle applicazioni tecnologiche ed è all'origine della «teoria dei sistemi» e, in parte, della cibernetica. In queste teorie un sistema viene considerato euristicamente come «una struttura nella quale si può immettere qualche cosa (materia, energia, informazione) in certi istanti e che esso stesso emette qualcosa in certi istanti...» [Kalman, Falb e Arbib 1969, p. 4]. Da questo punto di vista il sistema viene visto come una «scatola nera» che a un certo istante viene stimolato da qualcosa (input, o variabile di controllo) e reagisce con un responso (output).

L'unica cosa che interessa la teoria sono le relazioni funzionali (o statistiche) tra input e output, e solo nei termini in cui tali relazioni sono note la descrizione del sistema può dirsi «completa». La principale differenza con il meccanicismo sta nel fatto che in questo caso la «fotografia istantanea di un sistema», cioè il suo output in un dato istante, non contiene piú in sé l'embrione del suo divenire, poiché questo dipenderà in genere anche dagli input passati e da quelli futuri; piú precisamente si suppone che l'output ad ogni istante t sia determinato dagli input negli istanti precedenti a t (causalità). In quest'approccio il concetto di «stato di un sistema» nel senso meccanicistico scompare completamente. C'è un approccio piú generale, che include propriamente le teorie meccanicistiche, e che sintetizza i casi del sistema come «scatola nera» e del sistema isolato. In quest'approccio accanto agli input e agli output, che rappresentano l'interazione del sistema con il mondo esterno, si considera anche un complesso di grandezze osservabili intrinseche al sistema mediante le quali è possibile definire un concetto di «stato» (istantaneo) del sistema con la stessa procedura usata nelle teorie meccanicistiche. Se B è un insieme completo di osservabili intrinseche del sistema, lo «stato» del sistema in un dato istante è determinato dai valori di B in quell'istante e l'output in quell'istante è determinato come funzione dello stato («funzione di lettura»). Tuttavia lo stato in un

dato istante non determina univocamente il divenire del sistema; questo è determinato anche dall'insieme degli input ricevuti dal sistema fino a quell'istante, rappresentato, nel modello matematico, come una funzione (funzione input) del tempo a valori nello spazio degli input o dei controlli, nel senso espresso dal seguente postulato (che generalizza il postulato meccanicistico del determinismo):

POSTULATO DI CAUSALITÀ. *Se B è un insieme di osservabili i cui valori definiscono gli stati del sistema ed $s<t$ sono due istanti di tempo, allora per ogni funzione input ω esiste una funzione $T^B_{t,s}(\omega)$ tale che $B_t = T^B_{t,s}(\omega)\, B_s$.*

La funzione $T^B_{t,s}(\omega)$ definisce quindi l'evoluzione del sistema in funzione del «parametro di controllo» ω, e il postulato della causalità si esprime nell'affermazione che l'evoluzione del sistema dall'istante s a t dipende solo dagli input anteriori a t (il passato determina il presente); oppure, equivalentemente: se ω e ω' sono due funzioni input tali che $\omega(u) = \omega'(u)$ per $u \leq t$, allora si ha che $T^B_{t,s}(\omega)\, B_s = T^B_{t,s}(\omega')\, B_s$. La teoria delle catastrofi di Thom rientra in questa classe generale di teorie e corrisponde al caso in cui non ci sono output e il parametro di controllo ω è indipendente dal tempo (i «punti di catastrofe» corrispondono a quei valori del parametro di controllo ω in cui l'evoluzione $T^B_{t,s}(\omega)$ cambia il suo comportamento qualitativo). I sistemi descritti dalle consuete teorie meccanicistiche corrispondono al caso in cui lo spazio dei controlli è banale (ridotto a un solo punto). Anche i sistemi del tipo «scatola nera» rientrano in questa classe (se le funzioni input e output soddisfano certe condizioni di compatibilità), nel senso che per tali sistemi un teorema generale assicura l'esistenza di uno «spazio degli stati», una «funzione di lettura» e una famiglia di evoluzioni $T_{t,s}(\omega)$, in modo tale che l'output all'istante t sia proprio quello ottenuto applicando la «funzione di lettura» allo stato, all'istante t. Lo «spazio degli stati» definito da questo teorema, oltre a non essere univocamente definito dai dati, è un oggetto molto astratto e privo di significato fisico immediato.

Il «problema della realizzazione», che ha ricevuto considerevole attenzione nella teoria dei sistemi, consiste nel tentativo di definire in modo costruttivo, fisicamente significativo, e in qualche modo intrinseco (cioè unico a meno d'isomorfismi) uno spazio degli stati associato, nel senso spiegato sopra, a un dato sistema input-output. Anche nel caso di sistemi semplici (sistemi lineari) la soluzione di questo problema è abbastanza complessa, e non si entrerà qui nei dettagli tecnici. È interessante però osservare che, mentre nel caso delle teorie meccanicistiche è un fatto storico che l'evoluzione determini lo spazio degli stati (cfr. § 5), nel caso della teoria dei sistemi il problema di *costruire* uno spazio degli stati associato a una data corrispondenza input-output si pone fin dall'inizio in modo esplicito come problema teorico.

7. *Forma debole del principio d'indeterminazione.*

La radice della differenza tra i concetti di stato in una teoria meccanicistica e in teoria quantistica sta nel principio d'indeterminazione di Heisenberg. In questo si possono distinguere due aspetti: uno, qualitativo, che consiste nella negazione del postulato IV) delle teorie meccanicistiche (il postulato di compatibilità); e uno, che esprime in forma quantitativa il grado d'incompatibilità tra piú grandezze. Si comincerà con l'analizzare le conseguenze del primo aspetto di questo principio (forma debole del principio d'indeterminazione di Heisenberg), enunciandolo nella forma di

POSTULATO H. *Esistono osservabili che non possono essere misurate simultaneamente con precisione arbitraria.*

In particolare, dal postulato H segue che esistono grandezze i cui valori esatti non possono essere misurati simultaneamente. Per tali grandezze la questione se una sia o no funzione di un'altra, e perciò il postulato I) delle teorie meccanicistiche, è priva di senso da un punto di vista operativo, poiché la verifica dell'affermazione $A_s = f(B_s)$ presuppone che si disponga di una tabella su cui siano elencati tutti (o, per una realistica verifica parziale, almeno alcuni) valori simultanei delle grandezze A e B all'istante s, ed è proprio la possibilità di costruire una tale tabella mediante concrete operazioni fisiche, che viene negata dal principio di Heisenberg.

Una qualsiasi affermazione riguardante lo stato del sistema inteso come complesso di tutte le proprietà fisiche del sistema in quell'istante, in un contesto che accetti il postulato H, è un'affermazione non scientifica in senso popperiano, cioè non falsificabile sperimentalmente, neppure con un esperimento ideale. In un tale contesto perdono ogni significato operativo sia il postulato dell'accessibilità sia il postulato del realismo (nella formulazione data al § 2) e, con essi, la concezione stessa dello «stato fisico» cosí come essa emerge dai postulati delle teorie meccanicistiche, cioè come fotografia istantanea e completa del divenire di un sistema. A una tale concezione non corrisponde piú, neppure a livello virtuale, alcun concetto nel modello matematico. Ciò che segue dal principio d'indeterminazione non è la negazione del postulato del realismo e dell'associato concetto di stato fisico, ma solo la non-intrinsecità di questi concetti alla nostra rappresentazione quantitativa della realtà. Essi possono essere conservati, rispettivamente come convinzione filosofica e come base euristica dell'immagine intuitiva del concetto di stato, nell'ambito di quelle convinzioni meta-fisiche che, come Lakatos ha dimostrato, svolgono un ruolo molto importante nello sviluppo e nell'interpretazione di ogni teoria fisica pur non facendone propriamente parte. Come Laplace non include Dio, nella sua descrizione meccanicistica del sistema della natura, perché «non ha bisogno di questa ipotesi», la teoria quantistica fa a meno del concetto di stato come fotografia istantanea e completa del divenire di un sistema. Come Kant considera l'esistenza di Dio un postulato, non delle scienze naturali, ma della vita morale, cosí la teoria

quantistica considera questo concetto di stato, e piú in generale il postulato del realismo, come prodotti del nostro rapporto psicologico con gli oggetti piuttosto che conseguenze necessarie delle nostre esperienze. Questo è il punto di vista formalizzato nel postulato H, il quale non è una affermazione sulla natura in sé (come il postulato del realismo), ma sui limiti a priori di ogni teoria della natura in quanto teoria sperimentale. Tali limitazioni a priori riguardano la quantità d'informazione simultanea conseguibile su un sistema.

Il concetto di «stato quantistico» esprime quanta parte dell'informazione sullo stato – inteso nel senso intuitivo delle teorie meccanicistiche – rimane controllabile con l'esperienza una volta accettate le limitazioni a priori imposte dal principio di Heisenberg. Quanto segue spiega il senso preciso in cui ciò avviene.

8. *I postulati statistici della teoria quantistica.*

In questo paragrafo ci si limiterà alla teoria quantistica non relativistica e il termine 'stato' verrà usato nel senso, spiegato nel § 3, della meccanica hamiltoniana classica di un sistema con un numero finito di gradi di libertà, distinguendolo dal termine 'stato quantistico'. Due (o piú) osservabili saranno dette *compatibili* se possono essere misurate simultaneamente.

Solo se C è un insieme di osservabili compatibili ha senso parlare di «una funzione di C» (cfr. § 2). Piú in generale, solo se C è un insieme di osservabili compatibili ha senso fisico parlare della probabilità condizionata $P(A=a \mid C=c)$ che l'osservabile A assuma il valore a se è noto che C assume il valore c. La nozione di probabilità condizionata generalizza quella di funzione, a cui si riduce nel caso in cui le probabilità assumano solo valori banali (piú precisamente: A è funzione di C se e solo se per ogni valore a di A e c di C la probabilità condizionata $P(A=a \mid C=c)$ può valere solo zero o uno).

Nel caso classico (meccanicistico), tutte le osservabili sono compatibili e un insieme completo di osservabili è sostanzialmente caratterizzato dall'essere un insieme di osservabili indipendenti, massimale rispetto a questa proprietà (infatti è chiaro che la completezza, nel senso del § 2, implica questa proprietà; e il viceversa è garantito, almeno localmente e per dipendenze funzionali regolari, dal «teorema delle funzioni implicite» di Dini). Tenuto conto di ciò si può affermare che, nel caso classico, un insieme completo di osservabili è un insieme B di osservabili indipendenti, massimale rispetto a questa proprietà e tale che per ogni osservabile A le probabilità condizionate $P(A=a \mid B=b)$ assumano solo i valori zero e uno.

Nel caso quantistico, un insieme completo di osservabili compatibili è definito come un insieme C di osservabili compatibili e indipendenti, massimale rispetto a queste due proprietà, e tale che per ogni osservabile A le probabilità condizionate $P(A=a \mid C=c)$ siano ben definite. Tutte le affermazioni di una teoria meccanicistica classica sono riducibili ad affermazioni del tipo: se B è un insieme completo di osservabili compatibili, allora il valore dell'osservabile $A = f_A(B)$ all'istante $t>s$, se è noto che $B_s = b$, è dato da $A_t = f_A(T^B_{t,s} b)$.

Analogamente, tutte le affermazioni di una teoria quantistica sono riducibili ad affermazioni del tipo: se C è un insieme completo di osservabili compatibili, allora la probabilità che il valore dell'osservabile A all'istante $t>s$ sia a, se è noto che $C_s=c$, è data da $P(A_t=a\,|\,C_s=c)$.

Queste probabilità sono dette «probabilità di transizione» o, in alcuni casi, per sottolineare la soggiacente intuizione meccanica, «sezioni d'urto». È principalmente dalla misura di queste quantità che la teoria quantistica ricava le sue informazioni sperimentali sulla natura. (Per l'esattezza, poiché la teoria quantistica prevede l'esistenza di osservabili i cui valori non sono esattamente misurabili – neppure nell'idealizzazione del modello matematico – talvolta l'espressione $P(A=a\,|\,C=c)$ rappresenta una densità di probabilità, o addirittura una densità di probabilità relativa. Per semplificare l'esposizione, nel seguito ci si limiterà alla discussione del caso in cui queste espressioni rappresentano delle vere e proprie probabilità condizionate).

In analogia a quanto fatto per le teorie meccanicistiche, e al fine di rendere piú chiaro il paragone con queste ultime, si enunceranno ora, nella forma di postulati, alcune affermazioni comuni a tutte le teorie quantistiche:

POSTULATO $\mathrm{I_S}$). *Esiste un insieme completo di osservabili compatibili.*

POSTULATO $\mathrm{II_S}$). *Se C è un insieme di osservabili compatibili e A una osservabile tale che a un certo istante t la probabilità condizionata che $A_t=a$ se $C_t=c$ è $P(A=a\,|\,C=c)$, allora per ogni istante s la probabilità condizionata che $A_s=a$ se $C_s=c$ è anche $P(A=a\,|\,C=c)$.*

POSTULATO $\mathrm{IV_S}$). *È possibile, in linea di principio, misurare i valori esatti, a un dato istante, di un qualsiasi insieme di grandezze compatibili e in un dato istante una grandezza non può assumere due valori diversi.*

POSTULATO $\mathrm{V_S}$). *L'esistenza e la forma delle leggi caratteristiche di cui ai punti* $\mathrm{II_S}$) *e* $\mathrm{III_S}$) *(cfr. oltre) sono accessibili alla conoscenza umana mediante la ragione e gli esperimenti.*

Come si vede, i postulati $\mathrm{I_S}$), $\mathrm{II_S}$), $\mathrm{IV_S}$) e $\mathrm{V_S}$) sono gli analoghi dei corrispondenti postulati delle teorie meccanicistiche. Per quanto riguarda il postulato III) si verifica facilmente che esso equivale ad affermare che l'assegnazione di uno stato (meccanicistico) del sistema a un qualsiasi istante s determina completamente i valori di una qualsiasi altra osservabile in un qualsiasi altro istante $t>s$. Perciò, formulato in termini piú strettamente operazionali, il postulato meccanicista del determinismo assume la forma:

POSTULATO III). *Siano: B, un insieme completo di osservabili; $s<t$, due istanti di tempo; A, una qualsiasi osservabile; I, una moltitudine di sistemi identici tali che, all'istante s, $B=b_s$. Allora esiste un valore di A – $a=f_A(b_s;\,t)$ – che non dipende dall'insieme I, tale che, se all'istante t, viene misurato A su tutti i sistemi dell'insieme I, il risultato della misura è sempre lo stesso, e cioè $f_A(b_s;\,t)$.*

La teoria quantistica indebolisce il postulato meccanicistico del determinismo nel modo seguente:

POSTULATO $\mathrm{III_S}$). *Siano: C, un insieme completo di osservabili compatibili; $s<t$, due istanti di tempo; A, una qualsiasi osservabile; I, una moltitudine di sistemi identici tali che, all'istante s, $B=b_s$. Allora esiste una distribuzione di probabilità sui valori di A – $P_A(\cdot\,|\,b_s;\,t)$ – che non dipende dall'insieme I, tale che, se all'istante t viene misurato A su tutti i sistemi dell'insieme I, la frequenza relativa del risultato $A=a$ è (approssimativamente) uguale a $P_A(A=a\,|\,b_s;\,t)$.*

A differenza dei postulati I)-V) che caratterizzano le teorie meccanicistiche, i postulati $\mathrm{I_S}$), ..., $\mathrm{V_S}$) sono soddisfatti dalle teorie quantistiche, ma non le caratterizzano. Al contrario, essi definiscono la piú generale teoria statistico-deterministica. Per questo motivo, prima di analizzare nella sua specificità il concetto di stato quantistico, e al fine di far risaltare meglio questa specificità, sarà opportuno discutere il concetto di stato in una generale teoria statistica.

9. *Il concetto di stato in una teoria statistica.*

Per «teoria statistica» in generale, s'intende una qualsiasi teoria che soddisfi le condizioni espresse dai postulati $\mathrm{I_S}$), $\mathrm{IV_S}$), $\mathrm{V_S}$) del § 8 (si ricordi che nella definizione di insieme completo di osservabili compatibili è inclusa la condizione che le probabilità condizionate da valori assunti da tale insieme siano ben definite per ogni osservabile della teoria). Se anche il postulato $\mathrm{II_S}$) è soddisfatto, si dirà che la teoria statistica in questione è stazionaria, o che il sistema da esso descritto è temporalmente omogeneo. Se anche il postulato $\mathrm{III_S}$) è soddisfatto, si dirà che la teoria statistica è di tipo deterministico (o markoviano). Come per le teorie meccanicistiche, anche in questo caso è possibile formulare i postulati in modo «relativistico», cioè senza attribuire al tempo un ruolo privilegiato ma, per limiti di spazio, non si discuterà qui una tale formulazione.

In una teoria statistica assegnare uno stato di un sistema equivale ad assegnare, per ogni osservabile, la distribuzione di probabilità dei suoi valori condizionata dal fatto che sono noti i valori assunti da un insieme di osservabili compatibili. Se l'insieme di osservabili compatibili, che definisce il condizionamento, è completo, spesso si parla di «stato puro»; in caso contrario, di «miscela».

Il concetto probabilistico di condizionamento traduce, nel modello matematico, il concetto fisico di preparazione di un sistema relativamente a un dato insieme di osservabili (talvolta, nella letteratura fisica, il termine 'preparazione' viene usato come sinonimo di condizionamento su un insieme completo di osservabili, ma nel seguito ci si atterrà all'uso piú generale del termine introdotto sopra).

Se B è un qualsiasi insieme di osservabili compatibili e b un suo valore, si denoterà $P(\cdot\,|\,B=b)$ lo stato da esso definito. Per definizione, quindi, se A è un'osservabile, $P(A=a\,|\,B=b)$ rappresenta la probabilità condizionata che A valga a se $B=b$. Se B è un insieme completo di osservabili compatibili, il concetto di stato (statistico) determinato da $B=b$ generalizza il corrispondente concetto meccanicistico nello stesso senso in cui la nozione di probabilità condi-

zionata generalizza quella di funzione. Nel caso meccanicistico lo stato a un dato istante s è determinato dall'assegnazione dei valori assunti in quell'istante dalle osservabili B di un insieme completo. Tale assegnazione determina implicitamente lo stato. Affinché la determinazione sia esplicita occorre conoscere per ogni osservabile A della teoria la forma precisa della sua dipendenza funzionale da B, cioè la funzione f tale che $f(B)=A$. Come si è detto, tale funzione è completamente determinata dalle probabilità condizionate $P(A=a\,|\,B=b)$; infatti, nel caso meccanicistico, fissato b esiste un unico a per cui tale probabilità vale uno, e questo è caratterizzato dalla proprietà $a=f(b)$.

Nella pratica una misura non è mai ottenuta con un solo esperimento, e i dati sperimentali da cui si deduce la relazione teorica $A=f(B)$ sono proprio le probabilità condizionate $P(A=a\,|\,B=b)$ o, per meglio dire, il loro corrispondente sperimentale, cioè le frequenze relative.

La dipendenza funzionale di A da B si manifesta nel fatto che la distribuzione $P(A=a\,|\,B=b)$ è praticamente tutta concentrata intorno a un unico valore. Se ciò non accade ma le frequenze relative dei risultati continuano ad essere – a meno di piccole fluttuazioni intorno a determinati valori – stabili, si parla di dipendenza statistica. Riassumendo: sia in una teoria meccanicistica sia in una teoria statistica, uno stato è determinato dall'assegnazione dei valori simultanei di un insieme di osservabili compatibili. Ma, per la determinazione esplicita (sia pure virtuale) dello stato, in una teoria meccanicistica occorre dare la forma delle dipendenze funzionali, in una teoria statistica la forma delle dipendenze statistiche; e queste dipendenze sono, in entrambi i casi, determinate dall'assegnazione delle probabilità condizionate $P(A=a\,|\,B=b)$.

Le considerazioni precedenti mostrano che la nozione di dipendenza funzionale esatta come caso particolare di dipendenza statistica non è un artificio formale, ma rispecchia fedelmente la prassi sperimentale, in cui le probabilità condizionate (qui intese nel senso di frequenze relative condizionate) costituiscono un concetto primario rispetto alle dipendenze funzionali. In questo senso si può dire che le distribuzioni condizionate hanno un grado di «realtà obiettiva» maggiore di quello delle dipendenze funzionali anche nei casi tradizionali in cui soltanto quest'ultimo tipo di dipendenze viene considerato. Da questo punto di vista, paradossalmente l'introduzione delle probabilità condizionate invece delle dipendenze funzionali esatte non introduce un elemento di soggettività nella teoria, ma uno di obiettività. Essa inoltre permette di guardare alla transizione dalla concezione meccanicistica alla piú generale concezione statistica, non come a un brusco salto qualitativo, a una discontinuità, ma come a un processo continuo. Se infatti si pensa a una dipendenza funzionale esatta come a una distribuzione di probabilità «molto concentrata» intorno a un valore, si può usare una variante dell'*argumentum acervi* della scolastica per dimostrare come la separazione tra «dipendenza esatta» e «dipendenza statistica forte» sia molto piú labile di quanto non si possa essere abituati a credere: infatti, è possibile deformare poco alla volta, con continuità, una distribuzione molto concentrata fino a renderla diffusa su piú valori, e cioè propriamente statistica; poiché il processo è continuo, non ci può essere in generale una netta separazione tra i

concetti di dipendenza funzionale e di dipendenza statistica, cosí come non ci può essere tra i concetti di «mucchio di grano» e di «N chicchi di grano» ($N=1, 2, 3, \ldots$) (si osservi che, come nel caso del «mucchio», l'argomento precedente è valido anche per osservabili che assumono valori discreti, purché numerosi e «poco differenti» l'uno dall'altro).

Fin qui si è parlato di come viene assegnato uno stato in una teoria statistica. Ma che cosa *è* uno stato in una tale teoria? Cioè, che cosa s'intende in una tale teoria, quando si dice che «il sistema si trova in uno stato determinato dall'assegnazione dei valori $B=b$»? Si è visto che lo stato meccanicistico evoca l'immagine di una fotografia istantanea e completamente dettagliata di un sistema, dell'embrione del suo divenire. A quali intuizioni si riferisce questo concetto di stato «statistico», apparentemente cosí astratto?

Per chiarire questo punto si comincerà ad esaminare la questione nel caso di un esempio particolarmente significativo: la meccanica statistica (dell'equilibrio). L'idea intuitiva di «stato di equilibrio di un sistema termodinamico» consiste nel fatto che, col trascorrere del tempo, le «proprietà macroscopiche» del sistema si mantengono costanti. In generale, le proprietà fisiche di un sistema sono definite dal fatto che determinate grandezze osservabili assumono determinati valori; si chiameranno osservabili macroscopiche quelle che determinano le proprietà macroscopiche del sistema. Se si pensa a un sistema termodinamico come a un aggregato di moltissimi sottosistemi, euristicamente una osservabile macroscopica è una funzione degli stati meccanici di tutti i singoli sottosistemi, che dipende in modo abbastanza uniforme da questi stati (per esempio la somma delle energie cinetiche dei sottosistemi o una sua funzione).

La descrizione del sistema secondo la meccanica classica permette di associare a questo uno spazio degli stati (meccanici) e una evoluzione temporale; inoltre, poiché ogni osservabile macroscopica è una funzione dello stato meccanico del sistema, la conoscenza dello stato meccanico del sistema nell'istante s permette di calcolare con esattezza il valore dell'osservabile macroscopica in ogni istante $t>s$ e, in particolare, le medie temporali di queste osservabili.

In uno stato di equilibrio, per quanto detto sopra, i valori di queste osservabili macroscopiche devono mantenersi costanti nel tempo a meno di piccole fluttuazioni. Ciò implica, per continuità fisica, che le medie temporali dei valori di tali osservabili esistano e che ciascuna osservabile fluttui di poco nel tempo intorno al suo valor medio. Nella descrizione termodinamica del sistema in generale le uniche quantità che interessano sono quelle macroscopiche. Perciò, per quanto detto sopra, se sono note le medie temporali delle osservabili macroscopiche si può affermare di conoscere tutte le proprietà del sistema rilevanti da un punto di vista termodinamico, ed è a tale complesso di proprietà che ci si riferisce quando si parla di «stato termodinamico» (di equilibrio) di un sistema. Assegnare uno stato termodinamico di equilibrio significa quindi assegnare l'insieme dei valori di tutte le osservabili macroscopiche identificati, a meno di piccole fluttuazioni, alle medie temporali di queste ultime. Come nel caso dello stato meccanico (per il quale si rimanda al § 3), anche in questo caso la costruzione del modello matematico è resa possibile dalla riduzione della infinità at-

tuale delle proprietà che determinano lo stato a una infinità potenziale, nel senso che non si assegnano direttamente le medie temporali di tutte le osservabili macroscopiche ma un unico oggetto matematico e un algoritmo che, almeno in linea di principio, permette di calcolare ciascuna di tali medie. L'oggetto matematico associato allo stato termodinamico di equilibrio è una misura di probabilità e il passaggio dal concetto di stato al suo modello matematico è reso possibile dalla «ipotesi ergodica», il cui contenuto concettuale può essere formulato nel modo seguente: sullo spazio degli stati meccanici del sistema esiste una misura di probabilità P, dipendente dallo stato iniziale soltanto attraverso parametri macroscopici, tale che la media temporale di una qualunque osservabile macroscopica relativa a uno stato iniziale con gli stessi valori dei parametri macroscopici è uguale al valore medio di questa osservabile sullo spazio degli stati meccanici calcolato con la misura P (per l'esattezza, l'ipotesi ergodica include anche l'ipotesi che, in generale per sistemi di interesse termodinamico, l'energia sia l'unica quantità che si conserva nel tempo; ma nella presente discussione qualitativa non ci si servirà di questo fatto). La validità dell'ipotesi ergodica in una vasta classe di sistemi è oggi pienamente confermata sia dai successi empirici di questa ipotesi, sia da molti teoremi matematici (un nuovo ramo della matematica, la «teoria ergodica», è nato dai tentativi di dimostrare l'ipotesi ergodica), sia da esperimenti numerici fatti al calcolatore. Tuttavia la demarcazione dei limiti precisi di validità di quest'ipotesi costituisce ancora oggi un formidabile problema aperto per la matematica, ed è probabile che l'investigazione di questi limiti conduca a scoperte interessanti sia per la fisica sia per la matematica.

Entro i limiti di validità dell'ipotesi ergodica, la misura di probabilità P permette di calcolare i valori di tutte le osservabili macroscopiche (e quindi tutte le proprietà termodinamiche) allo stesso modo in cui la conoscenza dello stato meccanico permette di calcolare i valori di tutte le osservabili (e quindi tutte le proprietà) meccaniche. In questo senso si dice che la misura di probabilità P determina (spesso, per abuso di linguaggio, si dice che *è*) uno stato di equilibrio (termodinamico) del sistema.

La meccanica statistica che, nata dalle idee di Maxwell, Boltzmann e Gibbs, è uno dei più grandi trionfi della fisica dell'Ottocento, permette di dedurre la forma esplicita delle misure P, che sono dette «misure di Gibbs», e dimostra che ciascuna di queste misure è univocamente determinata dall'assegnazione della temperatura del sistema termodinamico; per questo motivo si denoterà $P(\cdot\,|\,T=\tau)$ la (unica) misura di Gibbs definita dal valore τ della temperatura T. Si osservi infine che poiché ogni $P(\cdot\,|\,T=\tau)$ è una misura di probabilità sullo spazio degli stati meccanici del sistema, e poiché ogni altra osservabile è esprimibile come funzione dello stato meccanico, la $P(\cdot\,|\,T=\tau)$ fornisce una distribuzione di probabilità sui valori di una qualsiasi osservabile.

L'esempio della meccanica statistica è importante poiché per la prima volta, in questo caso, il termine 'stato di un sistema' viene riferito a una misura di probabilità. In quest'uso, inoltre, emergono alcune caratteristiche generali che si riscontreranno in tutte le altre teorie statistiche, meccanica quantistica inclusa. Il riferimento, specificamente, è ai seguenti aspetti: uno «stato» di un sistema

secondo la teoria in questione è determinato dall'assegnazione di una probabilità condizionata; l'informazione contenuta in questa probabilità è massimale all'interno della teoria.

Siano ora B e C due complessi di grandezze, ciascuno dei quali determina uno stato con ogni insieme dei suoi valori. È lecito chiedersi qual è la probabilità che C sia uguale a c con la condizione che B sia uguale a b, cioè – come si dice talvolta identificando lo stato con la condizione che lo determina – qual è la probabilità che il sistema si trovi in uno stato (quello definito da $C=c$) se uno sa a priori che esso si trova in un altro stato (quello definito da $B=b$). La risposta è fornita dalla probabilità condizionata $P(C=c\,|\,B=b)$ e perciò, in particolare, se si è in una teoria esatta, tale probabilità non può valere che 1 – quando le condizioni $B=b$ e $C=c$ definiscono lo stesso stato – oppure 0, altrimenti (mutua esclusività degli stati meccanicistici). Ma se l'insieme B non è completo nel senso del meccanicismo (cioè non tutte le altre osservabili sono funzioni di B) e la condizione $B=b$ definisce uno stato nella teoria che si sta considerando, allora non c'è niente di paradossale nell'affermare che il sistema che si trova nello stato definito da $B=b$ possa, con probabilità non nulla, trovarsi anche in un altro stato. Infatti con ciò s'intende soltanto dire che, se si è preparato il sistema in modo che risulti con certezza $B=b$, allora c' è una probabilità non nulla che la misura di C possa dare il risultato c. Questa è appunto la situazione che si verifica in teoria quantistica.

Per concludere, infine, si osservi che l'analisi del concetto di stato in una teoria statistica permette di unificare, da un punto di vista formale, tutte le definizioni del termine 'stato', che s'incontrano nelle varie teorie, nel modo seguente: data una teoria – caratterizzata da un insieme $\mathcal{O}$ di osservabili e da un insieme $\mathcal{A}$ di affermazioni, uno stato relativo alla teoria $\{\mathcal{O}, \mathcal{A}\}$ è determinato dall'assegnazione dei valori di un insieme (completo) di osservabili compatibili mutualmente indipendenti. Se B e C sono due insiemi siffatti, si dirà che le condizioni $B=b$ e $C=c$ sono equivalenti dal punto di vista operativo se le probabilità condizionate da esse determinate *coincidono*, e cioè se per ogni osservabile A e un suo valore a, risulta che $P(A=a\,|\,B=b)=P(A=a\,|\,C=c)$. Non è difficile verificare che le condizioni $B=b$ e $C=c$ sono operativamente equivalenti se e solo se ogni volta che B assume il valore b, C assume il valore c e viceversa (in particolare i due insiemi di grandezze B e C sono l'uno funzione dell'altro, se e solo se per ogni valore b di B esiste un valore c di C tale che le due condizioni $B=b$ e $C=c$ siano operativamente equivalenti). Ciò posto, si può dire che uno stato relativo alla teoria $\{\mathcal{O}, \mathcal{A}\}$ è una classe di equivalenza di condizioni operativamente equivalenti. Nel caso delle teorie meccanicistiche una tale classe di equivalenza determina univocamente un punto nello spazio astratto degli stati e viceversa a ogni punto in questo spazio astratto corrisponde un'unica classe di equivalenza nel senso specificato sopra. Nel caso di una generale teoria statistica si può solo dire che l'assegnazione di uno stato equivale ad assegnare una distribuzione di probabilità sui valori di ciascuna osservabile della teoria. Tale distribuzione di probabilità sarà banale, cioè concentrata su un solo punto, per tutte quelle osservabili esprimibili come funzioni di uno qual-

siasi degli insiemi completi i cui valori determinano lo stato. Nel primo caso, il passaggio dallo stato astratto a una rappresentazione concreta si effettua scegliendo un sistema di coordinate (locali); nel secondo, scegliendo una particolare condizione $B=b$ tra tutte quelle ad essa operativamente equivalenti.

Mentre gli assiomi che definiscono le teorie meccanicistiche sono abbastanza forti da determinare completamente almeno il tipo qualitativo dello spazio degli stati (cfr. § 3), ciò non può accadere, proprio a causa della loro generalità, per gli assiomi delle teorie statistiche. In questo caso si avranno diversi modelli matematici a seconda delle ipotesi specifiche sulle osservabili e sulle probabilità condizionate. Si è già visto come i modelli delle teorie meccanicistiche, in quanto particolari teorie statistiche, sono inclusi in questa classe. Un altro esempio si ha quando tutte le osservabili della teoria sono esprimibili come funzioni di una certa famiglia di grandezze; in questi casi uno stato sarà determinato da un'unica misura di probabilità definita sullo spazio S dei valori di queste grandezze e perciò lo spazio degli stati sarà un insieme di misure di probabilità sullo spazio S (questo è il caso, per esempio, della meccanica statistica classica).

10. *Stati quantistici.*

La teoria quantistica accetta il principio d'indeterminazione di Heisenberg. Essa è quindi necessariamente una teoria statistica, nel senso che se due osservabili sono incompatibili, la conoscenza esatta del valore di una di esse in un dato istante implica che sui valori dell'altra la nostra informazione non potrà essere che di tipo statistico. Come in tutte le teorie statistiche, in meccanica quantistica uno stato è una famiglia di distribuzioni di probabilità (una per ogni osservabile) ed è completamente determinato dall'assegnazione dei valori di un insieme completo di osservabili compatibili.

Affermare che un sistema S si trova all'istante t in un dato stato quantistico fornisce un'informazione di tipo *esatto* su S, e cioè i valori assunti in quell'istante da un qualsiasi insieme di osservabili che definisce lo stato o da una loro funzione. Si tratta di una informazione assolutamente obiettiva che riguarda il singolo sistema S. Di tutte le osservabili, che non sono funzioni di quelle che definiscono lo stato, si conoscono solo le distribuzioni di probabilità in quell'istante. Anche questa è un'informazione obiettiva, poiché le frequenze relative previste dalla teoria quantistica si accordano bene con i dati sperimentali, e, in questo senso, un'affermazione del tipo «La probabilità che accada l'evento X è p» è altrettanto obiettiva quanto una del tipo «La lunghezza di questo tavolo è l». Tuttavia questa obiettività non riguarda piú il singolo sistema S ma, come accade per ogni teoria statistica, riguarda una (qualsiasi) moltitudine di sistemi identici a S e tutti i preparati in modo tale che all'istante t le osservabili che definiscono lo stato assumano gli stessi valori su ciascun sistema; una tale moltitudine spesso viene chiamata un *ensemble* statistico. Come si è osservato nel § 9, gli stati di una teoria statistica non sono in generale mutuamente esclusivi; nel caso particolare della teoria quantistica, il fatto che il sistema all'istante t si

trovi nello stato definito dalla condizione $B=b$ non esclude affatto che esso si trovi anche nello stato definito dalla condizione $C=c$, dove B e C sono due arbitrari insiemi completi di osservabili compatibili tali che B e C non siano tra loro compatibili e b, c due loro valori. Se, come implicitamente fanno molti autori, si associa al termine 'stato' una intuizione meccanicista e non probabilista, una simile situazione può apparire paradossale. In effetti, in un ensemble statistico definito dalla condizione $B=b$ la misura di B (o di una sua qualsiasi funzione) su uno qualunque dei sistemi dell'ensemble fornisce sempre b (o $f(b)$), mentre la misura di C darà luogo a risultati diversi di cui non è possibile prevedere nulla se non la frequenza relativa dei singoli risultati. In un ensemble statistico definito dalla condizione $C=c$ accade invece il contrario. Perciò, se si pensa alla proprietà «Essere nello stato definito dalla condizione $B=b$» come equivalente di «Appartenere ad un ensemble statistico definito dalla preparazione $B=b$», è chiaro che si arriva a una contraddizione con il fatto che il sistema può trovarsi nello stesso istante in due stati diversi. D'altra parte, dal punto di vista statistico, una moltitudine di sistemi preparati in modo tale che $B=b$ si comporta proprio come previsto da un ensemble definito dalla condizione $B=b$, e quindi per ciascun sistema di questa moltitudine l'esperienza sembrerebbe indicare la validità dell'identificazione, che la teoria ha appena dichiarato impossibile, tra «Essere nello stato definito da $B=b$» e «Appartenere ad un ensemble definito dalla condizione $B=b$».

Se si accetta questa identificazione, e quindi si ritiene che l'appartenere a un certo ensemble statistico sia una proprietà obiettiva della particella, allora l'atto della misura influisce direttamente su tale realtà. Per esempio, se il sistema all'istante s appartiene a un ensemble definito dalla condizione $B=b$ e all'istante $t>s$ su di esso si effettua una misura dell'insieme completo C di osservabili compatibili tra loro ma incompatibili con B, dopo la misura il sistema appartiene a un ensemble statistico definito dalla condizione $C=c$, e perciò se l'appartenenza a un ensemble è una «proprietà obiettiva» del sistema, si deve concludere che l'atto della misura ha cambiato una tale proprietà obiettiva.

Dal citato lavoro di Einstein, Podolsky e Rosen [1935] si può dedurre che tale conclusione contraddice il principio della relatività, secondo cui nessun segnale può propagarsi ad una velocità superiore a quella della luce. Alla radice di tali contraddizioni stanno, secondo una variante del ragionamento di Einstein, Podolsky e Rosen dovuta a Bohm, le seguenti considerazioni: si considerino due particelle appartenenti a un ensemble definito dalla condizione $B=b$; si supponga che l'insieme C consista di una sola osservabile che può assumere solo i valori $+1$ e -1 e che le due particelle siano vincolate in modo tale che se C vale $+1$ per la prima, C vale -1 per la seconda e viceversa (si può dimostrare che è possibile in natura preparare dei sistemi con queste proprietà). Si supponga inoltre che le due particelle siano molto lontane tra loro e che all'istante t si effettui una misura di C su una di esse con risultato, ad esempio, $C=+1$. Da quanto detto sopra, allora, a partire dall'istante t questa particella non appartiene piú all'ensemble statistico definito dalla condizione $B=b$, ma a uno definito dalla condizione $C=+1$; ma allora per le ipotesi formulate, a par-

tire dall'istante t l'altra particella appartiene all'ensemble statistico definito dalla condizione $C=-1$, e perciò, se il passaggio da un ensemble a un altro corrispondesse a un cambiamento di proprietà reali delle particelle, un'azione su una particella (la misura di C) avrebbe provocato il cambiamento istantaneo di una proprietà di un'altra particella che si trova a (grande) distanza dalla prima. Ma l'azione a distanza contraddice il principio della finitezza della velocità di propagazione delle interazioni.

La conclusione di Einstein, Podolsky e Rosen, i quali accettano il postulato della relatività, è che la seconda particella aveva la proprietà $C=-1$ anche prima che si effettuasse la misura di C sulla prima e che quindi la teoria quantistica è una teoria incompleta poiché essa, sulle proprietà $C=\pm 1$ per le due particelle sotto la condizione $B=b$, può fare soltanto previsioni statistiche. Altri autori, invece, ritengono (sulla base di argomentazioni tecniche che non verranno riportate qui) che il fatto che l'osservabile C assuma un valore ben preciso per ciascuna delle due particelle anche prima dell'operazione di misura è incompatibile con il formalismo matematico della teoria quantistica (cioè: una tale ipotesi implicherebbe che lo stato quantistico che descrive il sistema delle due particelle sia una «miscela», mentre il modello matematico prescrive che sia uno stato puro ottenuto come sovrapposizione dei due stati).

Se ciò fosse vero, allora sarebbe effettivamente necessario abbandonare il postulato del realismo (cfr. § 3) e la sua affermazione secondo cui tutte le proprietà di un sistema sono ben definite in ciascun istante e indipendenti dal fatto che su questo sistema si effettui o meno un'operazione di misura. Occorrerebbe, con Heisenberg, il quale ha però in seguito ritrattato questa sua affermazione, riesumare il concetto aristotelico di *potentia*, e affermare che le proprietà di un sistema, non compatibili con una preparazione data, esistono solo allo stato latente e che il nostro atto di misura determina il passaggio della potenza all'atto. Ma in tal caso, allora, si dovrebbe arrivare alla conclusione, attribuita da D'Espagnat alla scuola di Copenhagen, che «nessuna proprietà fisica dovrebbe mai essere attribuita a un sistema quantistico, con l'unica possibile eccezione di quelle molto particolari del tipo: se A è lo strumento di misura con cui il sistema sarà messo in interazione, le proprietà fisiche del sistema sono esclusivamente quelle per misurare le quali A è stato progettato» [1976, p. 340].

Affermazioni di questo genere sono davvero conseguenze logicamente necessarie dei principî fisici della teoria quantistica? Per rispondere a questa domanda è opportuno analizzare il problema della variazione nel tempo di uno stato quantistico. Per fissare le idee si supponga che all'istante s lo stato sia definito dalla condizione $B_s=b$ – dove B è un insieme completo di osservabili compatibili – e si denoti $P(\cdot\,|\,B_s=b)$ la famiglia di misure di probabilità associate a questa condizione. Per il postulato del determinismo statistico sono allora univocamente definite le probabilità $P(A_t=a\,|\,B_s=b)$ dove A è una qualsiasi osservabile e t è un istante di tempo successivo a s. Si denoti $P(\cdot\,|\,t;\ B_s=b)$ l'insieme di tutte queste probabilità (al variare di A). Si postula (e il formalismo matematico è coerente con tale richiesta) che questa famiglia di misure di probabilità definisca uno stato del sistema (nel senso del § 9), cioè che esista un insieme completo di osservabili compatibili C tali che $P(\cdot\,|\,t;\ B_s=b)=P(\cdot\,|\,C_t=c)$ – vale a dire tali che il condizionamento $B=b$ all'istante s implichi (in questo caso – data la reversibilità delle evoluzioni quantistiche – equivalga) la preparazione $C=c$ all'istante t. La legge di evoluzione della teoria quantistica fornisce la relazione tra gli stati $P(\cdot\,|\,B_s=b)$ e $P(\cdot\,|\,C_t=c)$.

In generale quindi tale legge non dice come variano nel tempo i valori di certe grandezze, ma come variano nel tempo le implicazioni deducibili dal fatto che all'istante s si è acquisita una certa informazione sul sistema effettuando un insieme completo di misure compatibili. Si supponga ora di effettuare, all'istante t, una misura simultanea su un altro insieme completo di osservabili compatibili D. A questo punto si devono distinguere due casi: 1) le osservabili D sono compatibili con le osservabili C; 2) le osservabili D non sono compatibili con le osservabili C. Nel caso 1) poiché sia C sia D sono insiemi completi, dovranno essere l'uno funzione dell'altro e perciò la misura di D non fornisce alcuna nuova informazione sul sistema, se è nota la relazione funzionale che permette di esprimere D in funzione di C; in questo caso quindi lo stato del sistema all'istante t sarà ancora descritto dalla misura di probabilità $P(\cdot\,|\,C_t=c)$. Nel caso 2) invece la misura di D all'istante t distrugge l'informazione sui valori di C all'istante t poiché, secondo le analisi di Heisenberg e Bohr, tale misura introduce un disturbo sui valori di C che è dell'ordine di grandezza del quanto d'azione (e quindi non trascurabile se il sistema è microscopico) e la cui influenza sui valori di C non è sottoponibile a controllo sperimentale. In questo caso la preparazione del sistema, cioè la quantità massimale d'informazione conseguibile su di esso in un istante mediante una misura, cambia e lo stato del sistema all'istante t è definito dalla probabilità $P(\cdot\,|\,D_t=d)$, dove d è il risultato della misura di D: lo stato del sistema è quindi cambiato per effetto della misura, nel senso che è cambiata la sua preparazione. Per comprendere bene il ruolo della misura nel cambiamento di stato quantistico, è utile considerare il caso di una misura «incompleta» (secondo una terminologia introdotta da Feyerabend [1957, p. 123]), cioè una misura di D che *non* è seguita dalla lettura del risultato. In questo caso la misura di D_t ha distrutto l'informazione che si aveva sui valori di C_t; d'altra parte, questa informazione non è stata sostituita dall'informazione sui valori di D. Perciò l'informazione disponibile all'istante t non consente di determinare uno stato puro relativo alla teoria quantistica (cioè determinato dai valori di un insieme completo di osservabili).

Tuttavia è ancora possibile determinare uno stato in senso lato, cioè una distribuzione di probabilità sui valori di ciascuna osservabile della teoria. Nel caso in cui D possa assumere solo un insieme discreto di valori, con semplici considerazioni sulle frequenze relative si può dedurre che se A è una qualsiasi osservabile, la probabilità che il valore di A in un istante u successivo a t sia a, sotto la condizione di «preparazione incompleta» descritta sopra, è data da

$$\sum_d P(A_u=a\,|\,D_t=d)\cdot P(D_t=d\,|\,B_s=b) \tag{3}$$

(dove la somma è intesa su tutti i valori di d).

Per chiarire meglio la peculiarità del caso quantistico, è utile considerare i risultati a cui si perviene effettuando le stesse operazioni (cioè misura completa e misura incompleta) su un sistema classico, in cui cioè tutte le osservabili sono compatibili. In questo caso, dopo una misura completa, la probabilità che $A_u = a$ diventa $P(A_u = a \mid [D_t = d] \wedge [C_t = c])$, vale a dire la nostra informazione, in generale, aumenta; dopo una misura incompleta, invece, dal teorema delle probabilità composte segue che la stessa probabilità è data da $P(A_u = a \mid C_t = c)$, cioè nel caso classico una misura incompleta non fornisce alcuna informazione, e quindi non altera la probabilità (il che è intuitivamente evidente, poiché l'unica informazione fornita da una misura incompleta nel caso quantistico è quella relativa alla esistenza del disturbo e il modello classico è costruito sull'idealizzazione di assenza totale di disturbo). I risultati di quest'analisi sono schematizzati nella tabella 1.

Gli esempi seguenti (uno riferito al caso classico e uno al caso quantistico) illustrano in due casi particolari la situazione generale descritta dalla tabella 1.

Caso classico: un'urna contiene n palline, di cui n_r rosse e n_b bianche; la probabilità di estrarre a caso una pallina rossa è n_r/n. Se estraggo a caso k palline e non le guardo, la probabilità che la $(k+1)$-esima pallina estratta sia rossa è ancora, come si verifica immediatamente, n_r/n. Se però le guardo e scopro che, delle k estratte, k_r sono rosse e k_b bianche, allora la probabilità che la $(k+1)$-esima pallina estratta sia rossa diventa $(n_r - k_r)/(n-k)$. Si veda il diagramma *a*) nella figura 1.

Caso quantistico: si considerano delle osservabili che, in una opportuna unità di misura, possono assumere solo i valori ± 1 (spin in una data direzione spaziale). Siano α, β, γ, tre direzioni spaziali, e S_α, S_β, S_γ indichino rispettivamente lo spin in queste direzioni. Inizialmente il sistema è preparato in modo che la probabilità che S_α sia $+1$ è $\cos^2(\beta\alpha/2)$, dove $(\beta\alpha)$ è l'angolo tra le direzioni β e α. Se si misura S_γ e non si legge il risultato, secondo la teoria quantistica la probabilità che S_α sia $+1$ diventa $\cos^2(\beta\gamma/2)\cdot\cos^2(\gamma\alpha/2)+\sin^2(\beta\gamma/2)\cdot\sin^2(\gamma\alpha/2)$. Se si legge il risultato della misura di S_γ e si trova $+1$, la probabilità che S_α sia $+1$ diventa $\cos^2(\gamma\alpha/2)$ (come è illustrato nel diagramma *b*) della figura 1).

Tabella 1.

Variazione della probabilità $P(A_u = a \mid B_s = b)$ conseguente alla misura di D all'istante $t (s < t < u)$.

	Caso classico	*Caso quantistico (D e C incompatibili)*
Preparazione $B_s = b$	$P\{A_u = a \mid B_s = b\}$	$P\{A_u = a \mid B_s = b\}$
Misura di D_t senza lettura del risultato	$\sum_c P\{A_u = a \mid [D_t = d] \wedge [B_s = b]\} \cdot P\{D_t = d \mid B_s = b\} = P\{A = a \mid B_s = b\}$	$\sum_c P\{A = a \mid D_t = d\} \cdot P\{D_t = d \mid B_s = b\} \neq P\{A = a \mid B_s = b\}$
Lettura del risultato $D_t = d$	$P\{A = a \mid [D_t = d] \wedge [B_s = b]\}$	$P\{A = a \mid D_t = d\}$

Nella letteratura fisica il variare della probabilità dell'evento $A_u = a$ descritto nella tabella 1 (e riferito al caso quantistico) viene chiamato «collasso (o riduzione) del pacchetto d'onda». Il termine 'collasso' sta ad evocare il fatto che, al variare della condizione, la probabilità condizionata varia in modo brusco, discontinuo e immediato. Piú precisamente, in teoria quantistica il termine 'collasso del pacchetto d'onda' è, in generale, riferito non alla probabilità, ma alla densità di probabilità (o meglio al suo analogo quantistico: la matrice densità).

Fin qui è stato analizzato il modo in cui cambia lo stato quando si cambia la preparazione effettuando una misura a un dato istante. Nel seguito si analizzerà invece come cambia lo stato quando non si effettua alcuna misura; questo tipo di evoluzione descrive in che modo le implicazioni, deducibili su una misura al tempo t, da una preparazione fatta in un istante precedente s, variano con t. Si cominci a considerare il problema nel caso classico. In questo caso, la condizione del determinismo statistico si scrive nella forma

$$P(A_t = a \mid [C_s = c] \wedge [B_r = b]) = P(A_t = a \mid C_s = c) \tag{4}$$

per ogni $r < s < t$, e per ogni terna di osservabili A, B, C. Relazioni del tipo (4) sono chiamate, in teoria delle probabilità, «proprietà di Markov».

Dal teorema delle probabilità composte e dalla proprietà di Markov segue che

$$P(A_t = a \mid B_r = b) = \sum_c P(A_t = a \mid C_s = c) \cdot P(C_s = c \mid B_r = b) \tag{5}$$

(dove la somma è estesa a tutti i valori c di C).

Se $A = B = C$, introducendo la notazione $P^B_{t,s}(b', b) = P(B_t = b' \mid B_s = b)$ e

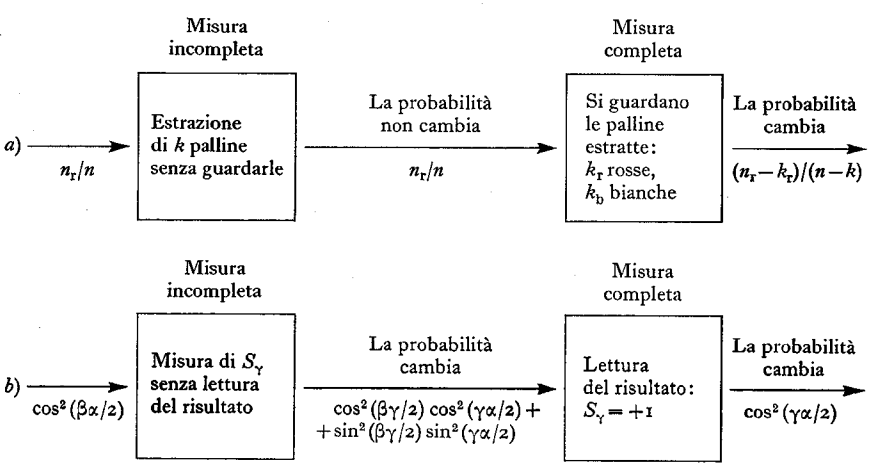


Figura 1.

Differenze tra le misure completa e incompleta nei sistemi classico e quantistico.

$P^B_{t,s} = (P^B_{t,s}(b',b))_{b',b}$ dalla equazione (5) si deduce immediatamente che le matrici che sono state indicate con $(P^B_{t,s})$ debbono soddisfare all'equazione

$$(6) \qquad P^B_{t,r} = P^B_{t,s} P^B_{s,r}$$

(in cui $r<s<t$) che, come si vede, è l'analogo dell'equazione (1), che descrive l'evoluzione di un sistema meccanicistico classico (in questo senso si parla di determinismo statistico). Anche in questo caso sotto alcune ipotesi tecniche si ricava dalla (6) la forma differenziale dell'equazione di evoluzione,

$$(7) \qquad \frac{d}{dt} P^B_{t,s} = X^B(t) P^B_{t,s},$$

dove, questa volta, $X^B(t)$ è una matrice.

Denotando $P(A_t|B_s)$, ..., le matrici con coefficienti $(P(A_t=a|B_s=b))_{a,b}$ dalle relazioni (5) e (7) segue che nel caso di sistemi temporalmente omogenei (quelli in cui le probabilità di transizione $P_{t,s}$ dipendono solo da $t-s$) si ha un'equazione di evoluzione del tipo

$$(8) \qquad \frac{d}{dt} P(A_t|B_s) = X^A P(A_t|B_s),$$

dove $X^A = X^A(0)$ soddisfa la (7) (con A al posto di B). Perciò, nel caso classico, se il postulato del determinismo statistico è soddisfatto, la forma matematica dell'equazione di evoluzione è data da un'equazione differenziale ordinaria del primo ordine nel tempo e lineare.

Nel caso della teoria quantistica, come si è già visto, la (5) in generale non vale, e perciò il ragionamento precedente non può essere applicato. In questo caso, l'evoluzione storica della teoria si è svolta secondo lo schema delineato al § 5, e cioè: *prima* si è scoperta una legge che permette di rendere conto in modo soddisfacente dell'evoluzione temporale delle quantità rilevanti della teoria (le probabilità condizionate), *poi* è stata la forma stessa di questa legge a suggerire un modello matematico per lo «spazio degli stati» di un sistema quantistico. Ma, al contrario di quanto accade per le teorie meccanicistiche, la nostra comprensione attuale della teoria quantistica non permette, neppure a posteriori e solo a livello qualitativo, di *dedurre* questo modello matematico da un complesso di principî di immediato contenuto fisico. Il modello matematico della teoria quantistica «funziona», è stato confermato da spettacolari successi, ma finora non è stato possibile chiarire in modo definitivo in che senso (e *se*) tale modello sia conseguenza necessaria e unica dei presupposti fisici della teoria. L'atteggiamento del mondo scientifico di fronte a questo problema non è uniforme. Alcuni ritengono che il successo del formalismo sia una indicazione della sua intrinsecità; altri ritengono possibile lo sviluppo di un formalismo alternativo e «fisicamente equivalente» (cioè un modello matematico diverso che conduce però alle stesse previsioni fisiche). Da entrambe le linee di pensiero sono emersi notevoli contributi alla nostra comprensione della teoria quantistica, ma per il momento il dibattito tra «il partito della razionalizzazione» e «il partito dell'alternativa» è da considerarsi ancora aperto. Per ora, non potendo *dedurre* il modello matematico dello spazio degli stati quantistico, ci si deve limitare a *descriverlo*.

Si è già visto che in una teoria statistica lo stato è definito dall'assegnazione di una distribuzione di probabilità su ciascuna osservabile della teoria, ed è determinato dall'assegnazione dei valori di un insieme completo di osservabili compatibili. Le prescrizioni della teoria quantistica per descrivere questa famiglia di misure di probabilità mediante un modello matematico unificato sono le seguenti (nelle quali, per semplicità, ci si limita a considerare osservabili che assumono un insieme discreto di valori):

1) Ad ogni condizione di preparazione $B=b$ si associa un proiettore (di rango 1) P^B_b definito su uno spazio di Hilbert complesso separabile H.
2) Dati due insiemi completi di osservabili compatibili A, B le probabilità condizionate corrispondenti sono date dalla formula

$$(9) \qquad P(A=a|B=b) = \tau(P^A_a P^B_b)$$

dove τ denota la traccia sugli operatori che agiscono in H.
3) A ogni osservabile A il cui insieme di valori è discreto, corrisponde nel modello matematico l'operatore $\bar{A} = \sum_a a \cdot P^A_a$, dove la somma è estesa a tutti i valori di A.
4) Le osservabili evolvono secondo un'equazione di evoluzione lineare

$$(10) \qquad \bar{A}_t = T_{t,s} \bar{A}_s$$

che definisce l'evoluzione delle probabilità condizionate mediante

$$P(A_t=a|B_s=b) = \tau(T_{t,s}(P^A_a) \cdot P^B_b).$$

Poiché, secondo quanto detto precedentemente, l'evoluzione deve trasformare condizioni di preparazione in condizioni di preparazione (cioè stati puri in stati puri), un teorema matematico dovuto a Wigner e a Kadison implica che l'evoluzione $T_{t,s}$ dev'essere indotta da operatori unitari.

In teoria quantistica, quindi, lo spazio degli stati di un sistema è l'insieme dei proiettori di rango 1 su uno spazio di Hilbert complesso separabile; spesso, identificando un vettore dello spazio di Hilbert al proiettore che esso definisce, si dice che lo spazio degli stati quantistici è uno spazio di Hilbert (o meglio, l'insieme delle rette passanti per l'origine di uno spazio di Hilbert). Le misure di probabilità associate a uno stato (puro) sono $P(\cdot|B=b) = \tau(P^B_b \cdot)$, e le «miscele» sono combinazioni convesse di stati puri. Come si vede, si tratta di un modello lontano dall'intuizione immediata, e lo è tanto piú in quanto finora non si è riusciti neppure a ridurlo a conseguenza necessaria di postulati «intuitivi».

Fin qui si è parlato del concetto di stato quantistico in un contesto non relativistico. La sintesi della relatività con la teoria quantistica, introducendo la possibilità che dalle masse si tramutino in energia o viceversa, oppure – con la scoperta dell'antimateria – la possibilità di creazione o annichilazione di coppie di particelle e antiparticelle – intacca profondamente il concetto di individualità

di un sistema e, a fortiori, il concetto di stato. In tale contesto, solo il campo conserva una sua individualità e le particelle vengono considerate come uno degli aspetti in cui può manifestarsi il campo. Si tratta però, ora, di un campo quantistico. Anche nel caso dei campi quantistici gli attuali modelli matematici dei corrispondenti spazi degli stati sono basati sul formalismo dello spazio di Hilbert, ma, nel caso dei campi quantistici, il modello matematico dello spazio degli stati si può dire compreso a fondo soltanto in casi particolari (campi liberi). Anche in questi casi la formazione di un adeguato concetto di stato è legata alla formulazione e alla soluzione, in modo soddisfacente sia dal punto di vista matematico sia da quello delle applicazioni, di un'appropriata legge di evoluzione e il tentativo di risolvere questo problema ha dato origine ad una delle linee di ricerca teoriche piú profonde e stimolanti della scienza attuale. [L. A.].

Einstein, A.; Podolsky, B.; e Rosen, N.
1935 *Can quantum-mechanical description of physical reality be considered complete?*, in «Physical Review», XLVII, pp. 777-80.

Espagnat, B. d'
1976 *Conceptual Foundations of Quantum Mechanics*, Benjamin, Reading Mass. 1976² (trad. it. Bibliopolis, Napoli 1980).

Feyerabend, P. K.
1957 *On the Quantum Theory of Measurement*, in S. Körner (a cura di), *Colston Papers*, IX. *Observation and Interpretation*, Butterworths, London.

Kalman, R. E.; Falb, P. L.; e Arbib, M. A.
1969 *Topics in Mathematical System Theory*, McGraw-Hill, New York.

Da un punto di vista euristico per 'stato' di un **sistema** fisico si intende un particolare modo di essere del sistema, piú specificamente un complesso di caratteristiche osservabili (cfr. **osservazione**) che ne determinino il comportamento (cfr. **comportamento e condizionamento**). È dunque possibile, nella pluralità delle accezioni del termine, isolare alcuni tratti comuni che permettono di unificare (cfr. **uno/molti**) le varie definizioni: ma l'unificazione a livello di definizione formale (cfr. **formalizzazione** e anche **logica**), che comporta una componente di carattere convenzionale (cfr. **convenzione**), non riflette necessariamente una **unità** delle intuizioni associate al termine, che dipendono invece dalle concezioni via via affermatesi di cosa significa descrivere un sistema nel senso di una **scienza** sperimentale (cfr. **empiria/esperienza, esperimento**). Gli sviluppi della meccanica statistica (cfr. **conservazione/invarianza, entropia, equilibrio/squilibrio, reversibilità/irreversibilità**) hanno indicato già come sul finire dell'Ottocento l'espressione 'stato di un sistema' vada riferita a una misura di **probabilità** (cfr. **induzione statistica**). Nel contesto quantistico (cfr. **quanti**) la concettualizzazione di cosa sia lo stato di un sistema cambia radicalmente le nostre categorie fisiche (cfr. **categorie/categorizzazione, concetto**) – come quelle di **caso/probabilità, ordine/disordine**, ecc. –, la sintesi di tale punto di vista con quello della **relatività** (introducendo l'eventualità che delle masse (cfr. **materia**) si tramutino in **energia** e viceversa, e la possibilità stessa di creazione o annichilazione di coppie di particelle e antiparticelle (cfr. **particella** e per altri aspetti **simmetria**)) intacca il concetto di individualità di un sistema, e con esso il concetto di stato (cfr. **forza/campo, interazione**).